BOKU - UNIVERSITY of
NATURAL RESOURCES and APPLIED LIFE SCIENCES
Institute of
HYDROBIOLOGY and AQUATIC ECOSYSTEM MANAGEMENT
Department of WATER, ATMOSPHERE and ENVIRONMENT

Classification and Analyses of Human Pressures in European Rivers

Diploma Thesis

Head of Institute:
O. Univ. Prof. Dr. Mathias JUNGWIRTH

Supervisor:
Ao. Univ. Prof. Dipl. - Ing. Dr. Stefan SCHMUTZ
Dipl. - Ing. Andreas MELCHER
Dipl. - Ing. Rafaela SCHINEGGER

Student:
Michael Rohrer

Vienna, March 2008

Acknowledgement

I would like to express my gratitude and thanks to my thesis supervisor Stefan Schmutz, Ao.Univ. Prof. Dipl.-Ing. Dr. for his overall help, valuable suggestions and patience while leading my diploma thesis. Many thanks belong also to Rafaela Schinegger, Dipl.-Ing. and Andreas Melcher, Dipl.-Ing. for their overall help and valuable suggestions during the correction of my thesis, and finally to Aneta Sinecká who gave me the force to study hard. Special thanks to my parents for their restless support.

Abstract

According to the Water Framework Directive (WFD) of the European Union (EU), which was adopted as a new legislation in the year 2000 the WFD restoration measures should lead to a “good status” of running waters by 2015 (WFD; 2000/60/EEC).

Therefore, new instruments are needed to optimize restoration measures. Particularly the WFD demands for economic analyses to evaluate the most efficient combination of measures.

One of the ecological instruments is the European Fish Index (EFI+) developed as a research project to gain new knowledge and to further develop and improve new biological assessment methods to meet needs of the Water Framework Directive (WFD). The output of the project will be a methodological approach to assess the ecological status of rivers in accordance with the WFD.

This diploma thesis focuses on the data mining procedures and the pressure index development within the ongoing EFI+ project. The first step of the thesis is to analyse the existing data from 15 European countries, which take part in the project, about the plausibility and the completeness. The next step are statistical analysis of the various pressure and their correlations. Different scenarios of indexes and data variants were developed and compared. This development should finally lead to a proposal for the development of a pressure index.

For the index development five scenarios with different completeness of data and different amount and type of pressure variables were developed (13 085 sites, max data loss (6 871 sites), consensus 1 (8 636 sites), consensus 2 (9 173 sites) and minimum data loss (9 498 sites). These five scenarios were compared again with three index approaches - the so-called arithmetic mean approach, degraded approach and worst-case approach.

Finally, the scenario “degraded approach with minimum data loss” was chosen as the most suitable for further application of pressure index and fish index development, because of the most adequate balanced distribution of the pressure status and completeness of the data.

Keywords: Water framework directive (WFD), European rivers, European fish index (EFI+), data mining, human pressures, pressure index, river types

Zusammenfassung

Die europäische Wasserrahmenrichtlinie (WRRL), welche im Jahr 2000 in Kraft trat, besagt, dass alle Gewässer bis ins Jahr 2015 in den „guten ökologischen Zustand“ überführt werden sollen. (RL 2000/60/EG; WRRL). Dafür werden von der Wasserrahmenrichtlinie Methoden zur Bewertung des ökologischen Zustandes von Fließgewässern gefordert.

Das von der EU im 6. Rahmenprogramm geförderte Projekt „Improvement and Spatial extension of the European Fish Index (EFI+) ist ein Forschungsprojekt, welches die Aufgabe hat mit Hilfe von Fischen, neue Methoden zur Bewertung des ökologischen Zustandes weiterzuentwickeln. Das EFI+ Projekt soll damit innerhalb der EU eine Harmonisierung und Standardisierung von fischökologischen Bewertungsmethoden herbeiführen.

Diese Diplomarbeit hat die Aufgabe mittels gezielter Datenanlyse anthropogene Belastungen im Rahmen des EFI+ Projektes in unterschiedlicher Art und Weise darzustellen.

Im ersten Schritt wurden alle Daten von 15 Europäischen Partnerländern dieses Projektes auf ihre Vollständigkeit und Plausibilität überprüft. Der nächste Schritt war die statistische Analyse der anthropogenen Einflüsse selbst und deren natürlicher Variabilität (Typisierung von Fließgewässern). Danach wurden verschiedene Szenarien und Varianten entwickelt, um Vergleiche und Vorschläge für die Entwicklung eines Belastungs-Index geben zu können.

Schließlich wurden fünf Szenarien mit verschiedener Vollständigkeit der Daten und verschiedenen Einflussvariablen verwendet. Je nach Vollständigkeit der Daten wurden 13085, 9498, 9173, 8636 und 6871 Erprobungsstellen bearbeitet.

Danach wurden die fünf Szenarien mit drei verschiedenen Index Varianten verglichen (Arithmetisches Mittel, Variante der mittleren anthropogenen Belastung von Fliessgewässern und Variante des schlechtest möglichsten Falles).

Am Ende konnte das Szenario mittlere anthropogene Belastung mit minimalem Datenverlust für die weitere Index Entwicklung ausgewählt werden, da es über den am plausibelsten und vollständigsten Datensatz für eine Weiterentwicklung des EFI verfügt.

Keywords:

Wasserrahmenrichtlinie, Europäische Flüsse, Europäischer Fischindex, Datamining, Anthropogene Einflüsse auf Flüsse, Flusstypologie

Contents

List of Figures

List of Tables

1. Introduction

According to the Water Framework Directive (WFD) of the European Union (EU), which was adopted as a new legislation in the year 2000 the WFD restoration measures should lead to a “good status” of running waters by 2015 (WFD, 2000). Therefore, new instruments are needed to optimize restoration measures. Particularly the WFD demands for economic analyses to evaluate the most efficient combination of measures

The aim of the Water Framework Directive is to create a European framework for the protection of inland surface waters, transitional waters, coastal waters and ground waters. This new legislation, now implemented in 25 EU member countries, strives for good ecological conditions in all surface waters.

In the past years river restoration has proceeded from actions of ameliorating impacts at the reach scale to serious plans to re-regulate or un-regulate entire catchments of large rivers, expressly to enhance natural attributes that have been measurably degraded (Stanford & Ward, 2001).

The WFD prescribes the following steps for ecological status assessment:

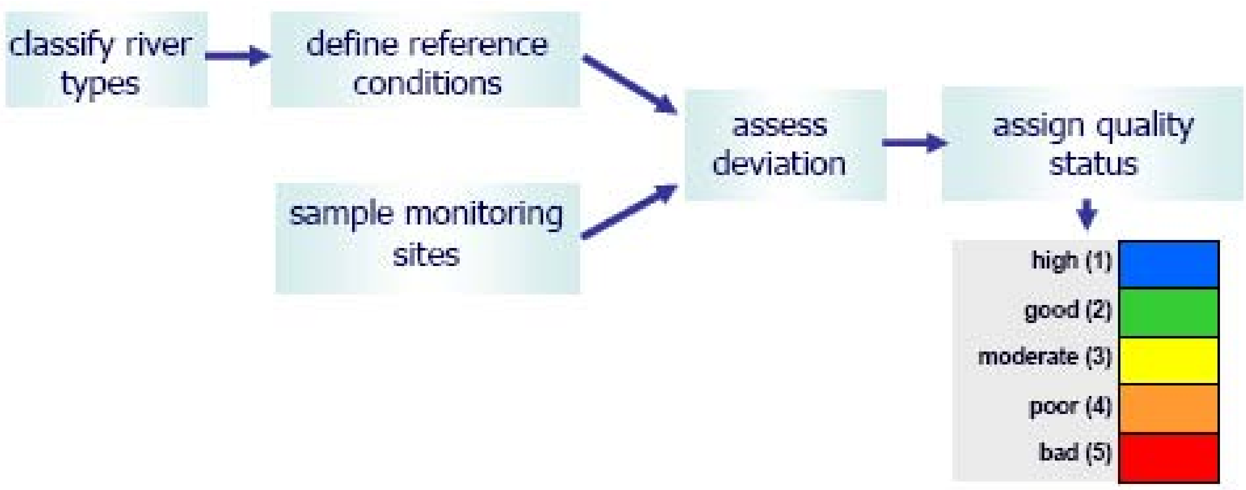

Figure 1: Steps prescribed in the WFD for ecological status assessment

Fish are beside the other three biological elements like (1) benthic invertebrate fauna (2) phytoplankton, (3) phytobenthos and macrophytes, one of four organism groups that can be used as an indicator to describe the ecological status of running waters. Fishes are, for the first time, part of a European monitoring network designed to observe the ecological status of running waters.

In the past different fish-based methods were used in Europe. Thus, the successful implementation of the WFD depends on the provision of reliable and standardised assessment tools. This was the motivation for the EC-funded research project FAME ”Development, Evaluation and Implementation of a standardised Fish-based Assessment Method for the Ecological Status of European Rivers”. The project aimed to develop, evaluate and implement a fish-based assessment method for the ecological status of European rivers to guarantee coherent and standardised monitoring throughout Europe (FAME 2005). The project time of FAME was 4 years, started in 2001 and was finished in 2004.

In order to further development of a standardized fish-based assessment method, the EC-funded research project “European Fish Index” (EFI+) started in 2007. FAME and EFI is based on the concept of the Index of Biotic Integrity (Karr 1981).

The principle of the Index of Biotic Integrity is based on the fact that fish communities respond to human alterations of aquatic ecosystems in a predictable and quantifiable manner. This Index is a tool to quantify human pressures by analysing alterations of the structure of fish communities. Karrs IBI used several components of fish communities, e.g. taxonomic composition, trophic levels, abundance and fish health. Each component is quantified by metrics (e.g. proportion of intolerant species). A metric is a measurable variable or process that represents an aspect of the biological structure, function, or other component of the fish community and changes in value along a gradient of human influence. Depending on the underlying biological hypotheses, a metric may decrease (e.g. number of sensitive species) or increase (e.g. number of tolerant species) with the intensity of human disturbances.

Fishes have proved their suitability as biological indicators for human disturbances for many reasons:

● Fishes are present in most surface waters.

● The identification of fishes is relatively easy and their taxonomy, ecological

requirements and life histories are generally better known than in other species groups.

● Fishes have evolved complex migration patterns making them sensitive to continuum

interruptions.

● The longevity of many fish species enables assessments to be sensitive to disturbance over relatively long time scales.

● The natural history and sensitivity to disturbances are well documented for many

species and their responses to environmental stressors are often known.

● Fishes generally occupy high trophic levels, and thus integrate conditions of lower trophic levels. In addition, different fish species represent distinct trophic levels: omnivores, herbivores, insectivores, planktivores and piscivores.

● Fishes occupy a variety of habitats in rivers: benthic, pelagic, rheophilic, limnophilic, etc., Species have specific habitat requirements and thus exhibit predictable responses to human induced habitat alterations.

● Depressed growth and recruitment are easily assessed and reflect stress.

● Fishes are valuable economic resources and are of public concern. Using fishes as indicators confers an easy and intuitive understanding of cause effect relationships to stakeholders beyond the scientific community.

The “FAME” project stands for “Development, Evaluation and Implementation of Standardized Fish-Based Method for the Ecological Status of Europeans Rivers in Europe”. The following 12 countries participated in the project: Austria, Belgium, France, Germany, Greece, Lithuania, Poland, Portugal, Spain, Sweden, the Netherlands and the United Kingdom.

On the 1 January 2007 the EU founded research project “Improvement and Spatial Extension of the European Fish Index (EFI+)” started. The project was designed to gain new knowledge, and further to develop and improve biological assessment methods, which should meet the standard needs of the Water Framework Directive (WFD).

The output of the project will be a standard methodological approach to assess the ecological status of the rivers of all members of the project in accordance with the WFD. Thus the EFI+ project represents a contribution to the Water Framework Directive for the further development and implementation of a fish-based assessment tools. This standard method can be used furthermore not only in all EU Member States, butalso in Candidate countries.

The FAME project, developed for Western and Northern Europe, was calibrated against an expert judgment based pre-classification of human pressure status. To predict biological reference conditions and quantifying the deviation of fish community structure from reference conditions on a statistical basis the project employs a number of environmental specialists.

Very large rivers under presented in a large context in the FAME project, but a wide range of river types was included in the development of the EFI+.

The overall objective of EFI+ project is to overcome existing limitations of the FAME project by developing a new, more accurate and pan-European fish index for all member states.

The scientific and technological objectives of the EFI+ project are:

• To evaluate the applicability of the existing EFI and the necessary improvements to the existing index in Central-Eastern Europe and Mediterranean ecoregions.

• To extend the scope of the existing EFI to cover very large rivers.

• To analyse relationships between hydro morphological pressures (incl. continuity disruptions) and fish assemblages to increase the accuracy of the EFI project.

• To adapt existing software to the requirements of the new EFI+ to allow calculation of the ecological status for running waters.

• And finally to implement and disseminate the EFI+ and supporting software by integrating of the project results into the CIS activities (Common Implementation Strategy) and ongoing national and international monitoring programmes such as the Joint Danube Survey.

The EFI+ team consists of 15 official partners countries.

The Netherlands and Lithuania are represented in the project by RIZA, who are being funded from national resources.

Furthermore, several partner institutions are integrated to the EFI+ group on a national level for data provision and expert support. Two external evaluators and an advisory group support the project team in achieving the scientific objectives and integrating end-users requirements properly.

2. Aim of the Thesis

The aim of this diploma thesis is to analyse different European anthropogenic pressures affecting the actual ecological status of European within the available data of the EFI+ project (http://efi-plus.boku.ac.at/), to give some proposal for a pressure index.

The specific objectives of the statistical analyses are as follows:

- To perform a comparison of pressure data on national level
- To analyse the natural variability of river types
- To analyse the pressure data about their plausibility and the completeness
- To analyse the correlation in between variables of pressure types
- To develop and analyse different methods to verify a pressure index
- To point out differences of the EFI+ and the FAME project
- To give some proposal for a pressure index

3. Basic tools - State of the Art

After more than two decades of study of the anthropogenic pressures on fish in rivers the facts and treats of this problem are well defined and can be found in several scientific limnology papers. On today's rivers, mostly not only one type of pressure at a time is influencing the in the water chemical, morphological or biological good status of rivers. Mostly, several of them can be found mixed together and furthermore the influence is rising with each pressure.

When an analysis of pressures affecting a river is done all pressures and their influence to each other should be taken in account. On the other hand, different kinds of anthropogenic pressures were often found to influence the same set of biotic characteristics of running waters.

Based on the experience of studies of the Institute of Hydrobiology and Ecosystem Management during several expert meetings and projects human pressures on rivers have been defined in the past (http://www.boku.ac.at/hfa/forschung/projekte.htm).

This knowledge was also used to define pressure criteria for the EFI+ project.

3.1. The FAME Project (Standardised Fish-based Assessment Method for the Ecological Status of European Rivers)

The FAME project, a project under the fifth R&D Framework Programme of the European Commission was founded from the idea to meet the required monitoring of riverine fish fauna for the Water Framework Directive (WFD) of the European Union because most of the EU member states did not have fish-based assessment methods compliant to WFD requirements. The main task of the project was to develop, evaluate and implement a standardised fish-based method for assessing the ecological status of European running waters. The following 12 countries participated in the project: Austria, Belgium, France, Germany, Greece, Lithuania, Poland, Portugal, Spain, Sweden, the Netherlands and the United Kingdom.

The FAME project developed and tested several fish-based assessment methods for the ecological status of rivers.

The European Fish Index (EFI) was selected as the method the most appropriate to satisfy the requirements of the water WFD.

The FIDES (Fish Database of European Streams) is the basic tool to develop the EFI (European Fish Index). The database is about 15 000 fish samples from 8 000 sites and 2 700 rivers of 16 European eco-regions contributed by 12 countries.

For each site, information about sampled fish, abiotic variables and human pressures were collected. The database also includes a wide list of European freshwater fish species assigned to ecological guilds according to their ecological characteristics.

The FIDES dataset contains different river types, reference sites and different levels of degradation of European rivers. Each fishing occasion was classified by the way of using a joint pressure variable to identify reference sites and levels of degradation.

For the calculation, morphological and hydrological conditions e.g.: nutrients (organic input), toxic substances (acidification) and connectivity were used.

The database of the FAME project consists of five main subject areas and five helping tables and the table metrics. The table site contains all characteristics of the sites. In the tables, catch lengths and length class, data of caught fishes presented. The Table Fishing occasion consists of: anthropogenic pressures, morphological variables and variables which are describing the sampling methods. In addition the database contains helping tables for historical data ecoregions, countries, reporters, and the table taxa and guilds. The table metrics is used for the in connection with the table fishing occasion for the index development. The figure 2 shows the structure of the FIDES database.

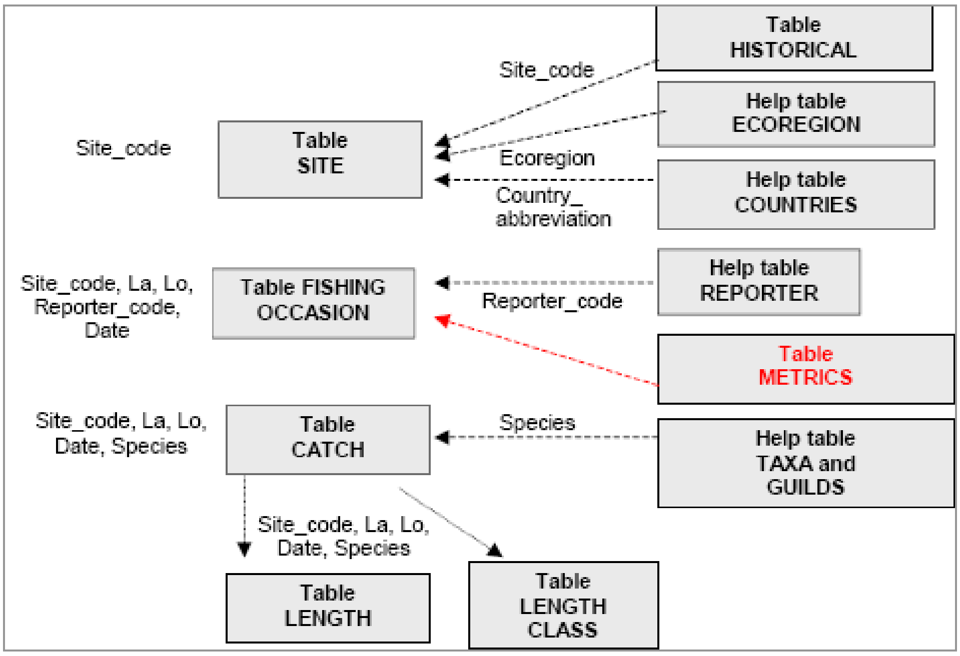

Figure 2: Structure of the central FIDES database, (FAME CONSORTIUM, 2004).

The figure 3 shows the general set up of the FIDES database by Hernandez (2003)

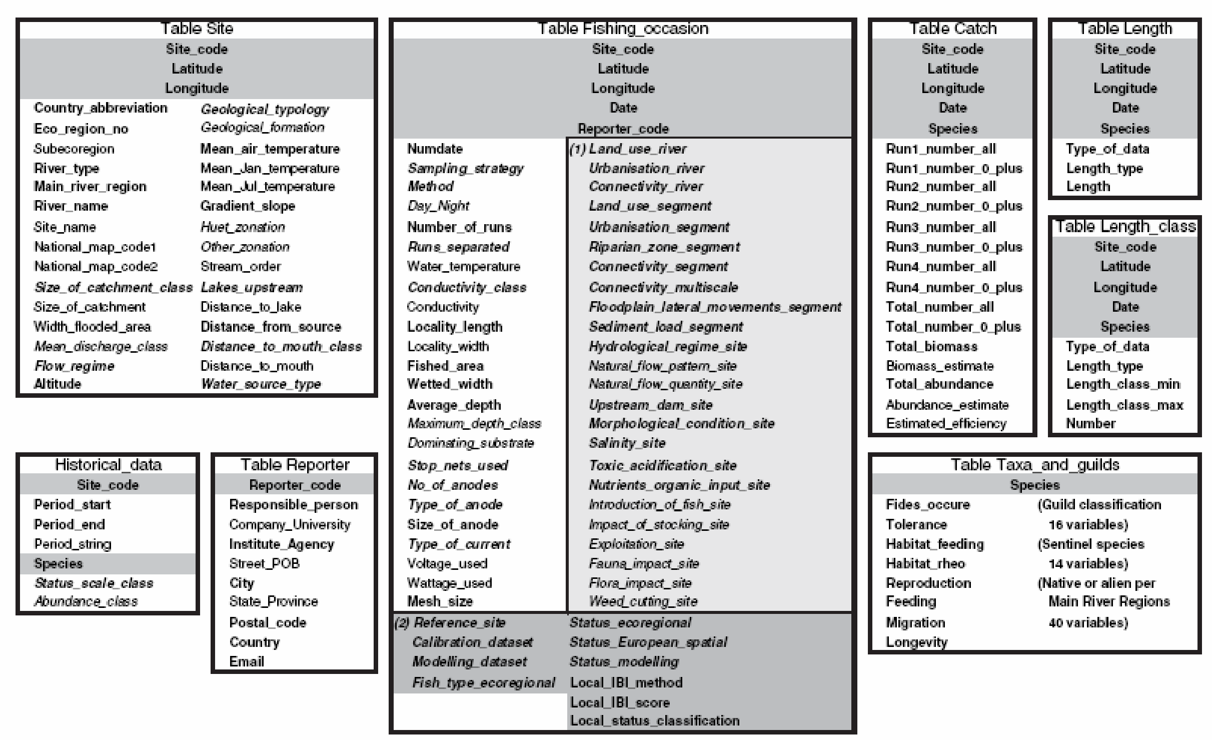

Figure 3: Structure of the FIDES database, (Hernandez 2003)

3.1.1. The Index Development of the FAME / EFI Index

The FAME project (http://fame.boku.ac.at/), a project under the fifth R&D Framework Programme of the European Commission was founded from the idea to meet the required monitoring of riverine fish fauna for the Water Framework Directive (WFD) of the European Union because most of the EU member states did not have fish-based assessment methods compliant to WFD requirements. The main task of the project was to develop, evaluate and implement a standardised fish-based method for assessing the ecological status of European running waters.

For the FAME project two different methodologies were used: the so-called spatially based modelling and the site-specific modelling, the latter leading to the European Fish Index (EFI). The advantage of the EFI is that, despite being a single index, it is applicable to a wide range of environmental conditions across Europe precluding the need for inter-calibration (Schmutz et al. 2007).

The EFI quantifies the deviation between predicted and observed conditions of the fish fauna derived from a predictive model that derives reference conditions for individual sites. The EFI consists of the following seven steps (figure 4):

1. An assessment metric was calculated from single-pass electric fishing catches. In the following ten metrics belonging to the following ecological functional groups were established: reproduction guilds, trophic structure, physical habitat, migratory behaviour and capacity to tolerate disturbance in general. Six of this metrics were based on species richness and four on were based on densities.

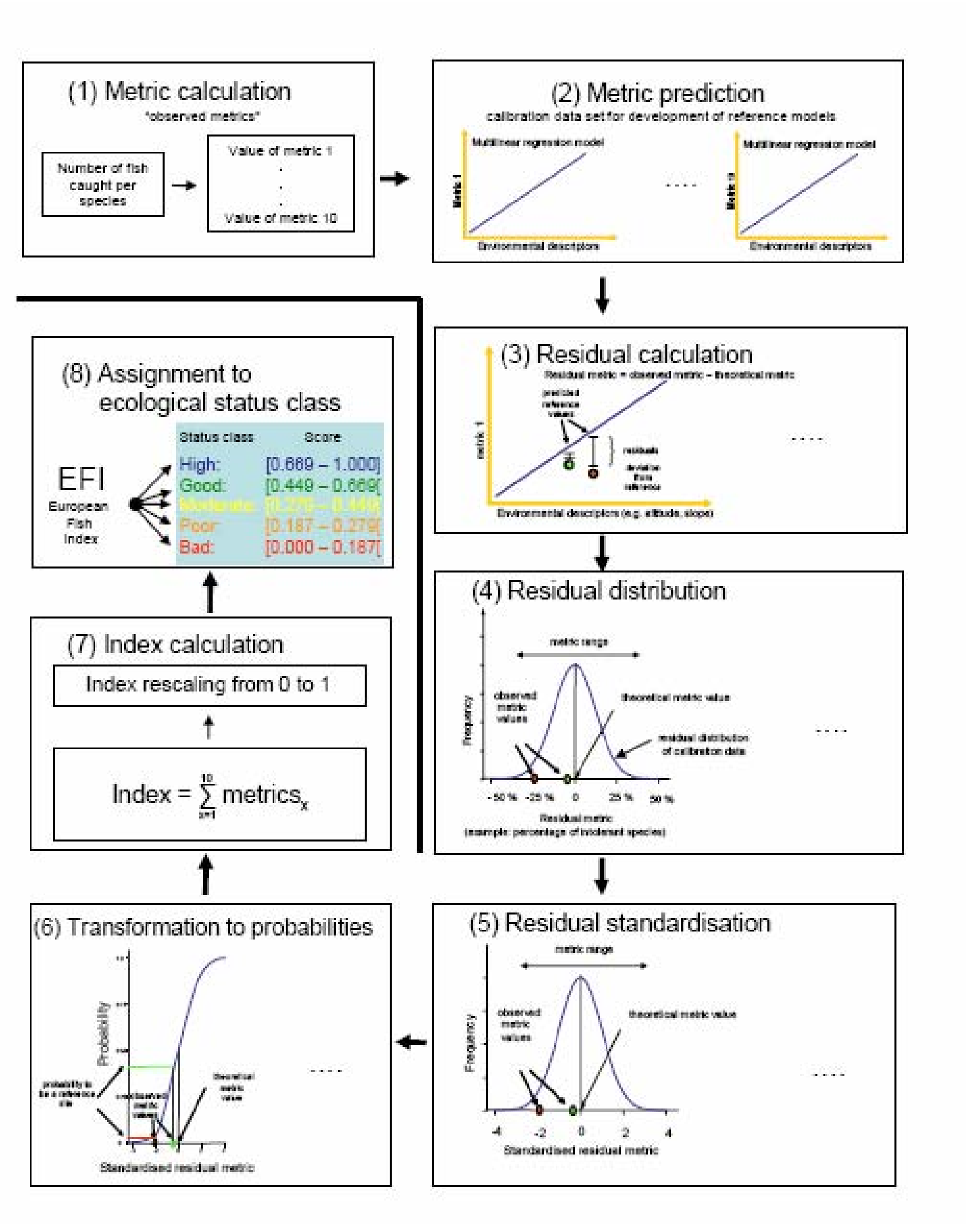

Figure 4: Methodology of the EFI index development (FAME CONSORTIUM, 2004).

2. For this step, a theoretical reference value describing the high or good status of a site (sites with only slight human disturbances) was evaluated for each metric that uses an environmental variable. This evaluation was done by a multilinear regression model calibrated with FIDES reference data.

3. To quantify the level of degradation the residuals of the multilinear regression models are used. These residuals are calculated as: observed values of the metric minus theoretical metric of the step number 2.

4. As the residual values scatter around the theoretical value, Impacted sites show greater difference from the theoretical value. Thus they are also less belong to the reference residual distribution than unimpacted or only slightly impacted sites.

5. As the metrics are based on different units it is necessary for a comparison to standardise them through a subtraction and a division by the mean and the standard deviation of the residuals of the reference sites.

6. At this step the residuals are transformed into probabilities. Because some standardised residuals values tend to increase with disturbance, whereas others decrease.

This transformation has two main advantages. Firstly all metrics will vary between 0 and 1, whereas the standardised residuals have no finite values. Secondly, all metrics will have the same response to disturbance, i.e. a decrease.

This final metric values of this step describe the probability for a site to be a reference site. Reference sites are defined as sites with the two best ecological integrity classes (1 and 2). The site that fits perfectly with the theoretical value will have a final metric value of 0.5. Values of impaired sites will decrease when the disturbance intensity increases. If for a site an probability higher than 0.5, is given, the in situ site is better than the predicted one. Also the probability for these sites that they belong to the integrity class 1 increases.

7. The final European Fish Index (EFI) is archived by summing the ten metrics, and then it is the score rescaled from 0 to 1.

3.2. The MIRR Project (A Model-Based Instrument for River Restoration)

The Model-based Instrument for River Restoration (MIRR) is thought to develop a strategic instrument for integrated assessment of restoration measures for running waters based on fish ecological criteria. As the Water Framework Directive (WFD) of the European Union demands that restoration measures should lead to the “good status” of running waters by 2015, he MIRR Project is an important input for the state of the art restoration methods (Schmutz et al. 2007). Download under: http://mirr.boku.ac.at/mirr_resultate.htm.

The project approach deals with the basic assumption, that the ecological status of running waters - when restored - is improving in a similar way, as it was deteriorated by pressures.

At the beginning of the project a list of pressure criteria's was developed by literature search and expert judgment for the following pressure types: canalization, continuum disruption, land use, impoundment, water abstraction and hydro peaking.

For the project a huge literature review of about 1150 collected papers, reports and reviews was done. These papers (N=331) contained relevant information to characterise the selected pressures. A list of significant criteria, regarding to the corresponding type of pressure was published. More than 350 potential criteria/features of running waters linking fish fauna and effects of human pressures were identified. Different kinds of human pressures were often found to influence the same set of abiotic characteristics of running waters.

The following table 1 shows a list of pressures and their keyword used in the MIRR-project.

Because pressure data wasn't available for whole Austria, the project took place in Lower Austria where the best data were available.

In total 938 samples taken from 715 sites are registered in the MIRR database. For 400 sites almost all pressure data regarding the pressure types: Channelisation, continuum disruption, land use, impoundment and water abstraction could be collected. In addition, also other environmental parameters were collected for the characterisation of the sites.

The Collected data were used to analyse the relationships between the reaction of fishes and pressures.

Table 1: List of pressures and keywords of the MIRR project (Zitek et al. 2006)

Channelisation	Continuum	Hydropeaking	Impoundment	Reservoir flushing	Water diversion
Degradation	Continuum	Hydro/power peaking	Impoundment	Reservoir/impoundment flushing/cleaning	Water diversion
Evaluation	Movement	Water-level fluctuations	Dam	Sediment(s) flushing/cleaning	Flow regulation
Habitat	Migration	Peak level	Reservoir	Reservoir desiltation	Water extraction
Impact	Floodplain	Abrupt, sudden, brusque discharge/water-level change	Barrier		Water abstraction
Incision	Connectivity	Sharp increase/decrease discharge/water-level	Weir		Water transfer
Regulation	Fragmentation	Flow fluctuations	River/stream regulation		Minimum flow
	Dam	Flood management			Residual flow
	Damm/Damming	Flow management			Environmental flow
	Hydroelectric				Flow regime alteration
	Tributary				

The MIRR Database

The MIRR database is based on the database of the FAME project and consists of 15 different subject areas. The subject of the tables site and fishing occasion contain all characteristics of the sites. In the tables catch and lengths the data of caught fishes are saved. All data regarding anthropogenic interferences are contained in the variables Morphology, Continuum, Impoundment, Land use, Residual flow, Water quality, index of human influence and Population. In addition the database contains subject areas of migration, historical data and fish metrics. For the calculation of the metrics the table Taxa and Guilds is necessary. The figure 5 shows the structure of the MIRR database.

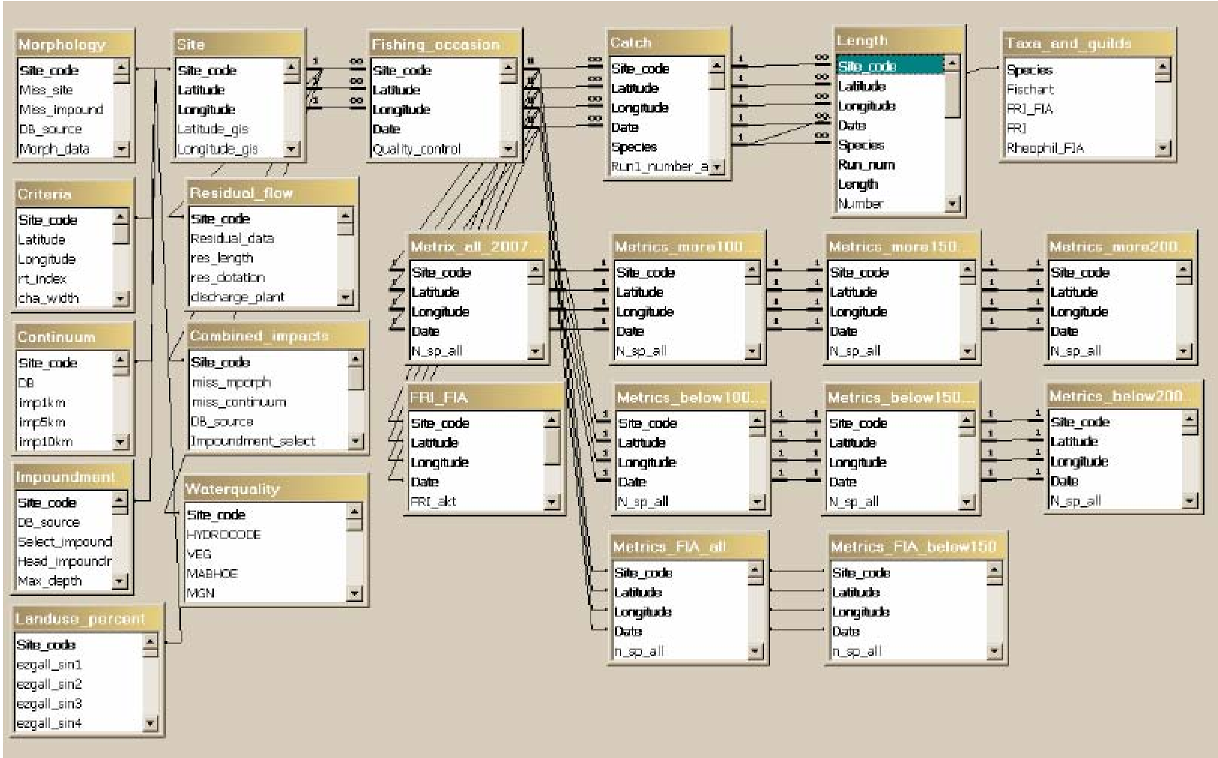

Figure 5: Structure of the database of the MIRR project, (Schmutz et al. 2007)

3.2.1. The Index Development of the (MIRR) Project: A Model-Based Instrument for River Restoration on Local Level

The Model-based Instrument for River Restoration (MIRR) was thought to develop a strategic instrument for integrated assessment of restoration measures for running waters based on fish ecological criteria.

For the evaluation of the fish ecological conditions, different pressure specific fish indexes were developed. This fish indexes consist of several fish ecological criteria's called "metrics". The development of this index follows the methodology of the "FAME" Consortium 2005, (Pont et al. 2006).

The index was calculated for five following pressure types (regulation of the river, river continuum, land use, residual flow and impoundments) This five pressure indices were calculated for rhithralic and potamalic zone. The output of this ten indexes was analysed by a discrimnat analyses to discover possible impacts. With this ten specific pressure indexes a combined index was developed after the principle "One out-all out" which is the same principle as the variant three "worst case "in this thesis. The index value with the highest divergence was at taken for the main fish index, which was used for further analyses.

3.3. The EFI+ Organisation, Time Table, Tasks and Working Programme

3.3.1. Aim of the EFI+ Project

The EFI+ project (http://efi-plus.boku.ac.at/), founded by the European Commission within the Sixth Framework Programme (2002-2006), is a research project for gaining new knowledge for the development and improvement of biological assessment methods that meet the needs of the Water Framework Directive (WFD).

The general aim of the EFI+ project is to develop for a standard methodological approach to assess the ecological status of rivers in accordance with the WFD.

The figure 6 shows the investigation area of the EFI+ Project. The following 15 countries participate in the project: Austria, Switzerland, Germany, Spain, Finland, France, Hungary, Italy, Lithuania, Netherlands, Poland, Portugal, Romania, Sweden, and the United Kingdom.

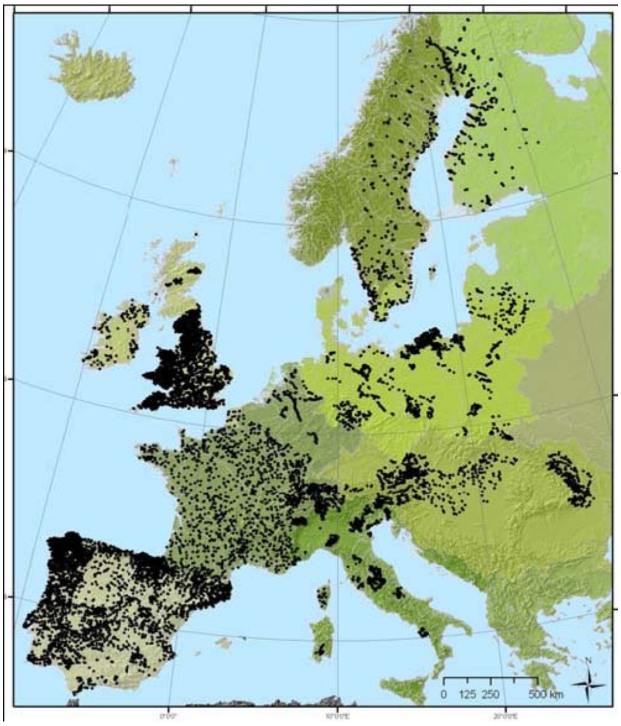

Figure 6: Spatial distribution of EFI+ sites

3.3.2. Basic Structure of the EFI+ Project

EFI+ consists of six technical work packages integrated through the overall project management. The project duration is 24 months, from the 1 January 2007 to the 31 December 2008. The table 2 gives an overview for the EFI+ project.

Table 2: The EFI+ project in an overview (EFI+ Newsletter March 2007)

Nr.	WP	Tasks	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
1	Basic tools	Fish data and species guilds	X	X																						
		Natural and anthropogenic descriptors	X	X																						
2	Data aquisition and management	Collecting fish data			X	X	X	X	X	X																
		Collecting natural and pressures descriptors			X	X	X	X	X	X																
		Database Management		X	X	X	X	X	X	X	X	X								X	X					
3	Pressures analysis and new metrics development	Pressure analyses								X	X	X	X	X												
		Evaluation of existing EFI									X	X	X													
		Hydromorphological pressures									X	X	X	X												
		Continuity disruption									X	X	X	X												
		Central/Eastern Rivers								X	X	X	X	X												
		Mediterranean Rivers								X	X	X	X	X												
		Large Floodplain Rivers								X	X	X	X	X												
		Low Species Rivers								X	X	X	X	X												
4	Fish Index Development	Modelling reference conditions					X	X	X	X	X	X	X	X	X	X	X									
		Metrics responses to pressures							X	X	X	X	X	X	X	X	X	X	X							
		Evaluation and sensitivity analyses																		X	X					
5	Implementation	Software and manual update										X									X	X	X			
		Catchment case-study (Danube survey)																				X	X			
6	Dissemination	End-user workshops						X																	X	X
		Web site	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
		Publications and external presentations													X	X	X	X	X	X	X	X	X	X	X	X
		International Conference										X						X						X	X	X
		Future exploitation of project results																		X	X	X	X	X	X	X
		Advisory Group	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X

The Water Framework Directive (WFD) monitoring programme was scheduled to be in operation for all member states of the European community by the end of the year 2006. The national data programme sampling will continue in year 2007 and in the following years. The result of the WFD monitoring programmes will be the basis for the river management basin plans. These plans should be ready by the year 2009.

The scheduled plan means that the EFI+ project, is appropriately able to deliver tools for assessing the field data collected in the monitoring programmes due to end in 2008.

Table 3: Timetable of the EFI+ project (EFI+ Newsletter March 2007)

Year-quarter	WFD requirements	*EFI*+ project schedule
2005-1		Characterisation of pressures
2006-4 / 2007-1	Monitoring programmes in operation	Presumed start of project
2007-2		Data collection
2007-4		Evaluation of existing EFI
2008-3		New version of EFI
2008-4	Draft river management basin plans	End of project
2009-4	River basin management plans	

During the initiation phase of the EFI+ project the basic principles and the variables for data collection were agreed. The focus in the EFI+ project lies on taken and now existing data from electro fishing surveys. The judgements about species composition in large rivers information were done also by other collected sampling methods.

The defined set of variables, which is provided for each sampling site and fishing occasion, enables the account for different features and fish amount of European rivers. Thus it defines adequately existing pressures at a sampling site. One of the biggest challenges for the selection of the variables was to find the best compromise between the time and financial resources available to the project.

Structure of the EFI+ Database

The figure 7 shows how the tables of the Access date base for the EFI+ are linked together. The database consists of 2 levels:

1. Level: Tables: Site, Fishing Occasion, Catch and Length

2. Level: Tables with Metadata, Fish guilds / taxa, Fish Owner Data, Diadromous, Reporter

Description of the Tables:

In each table the variable “sitecode” is the main linkages. As second linkage the variable “date” and others are used.

Table Site: information describing the site where the fish sample was taken

Table Fishing occasion: all anthropogenic pressures, morphological variables and variables which describe the sampling methods are described

Table Catch: all information of the catch procedure

Table Length: the total length of the fish is described

Table Diadromous: Historical data are described

Table Fishguilds/taxa: species names and family informations

Table Reporter: Information about the Institution who caught the fish on the site

Table: Fish owner: Information’s about the ownership of the river

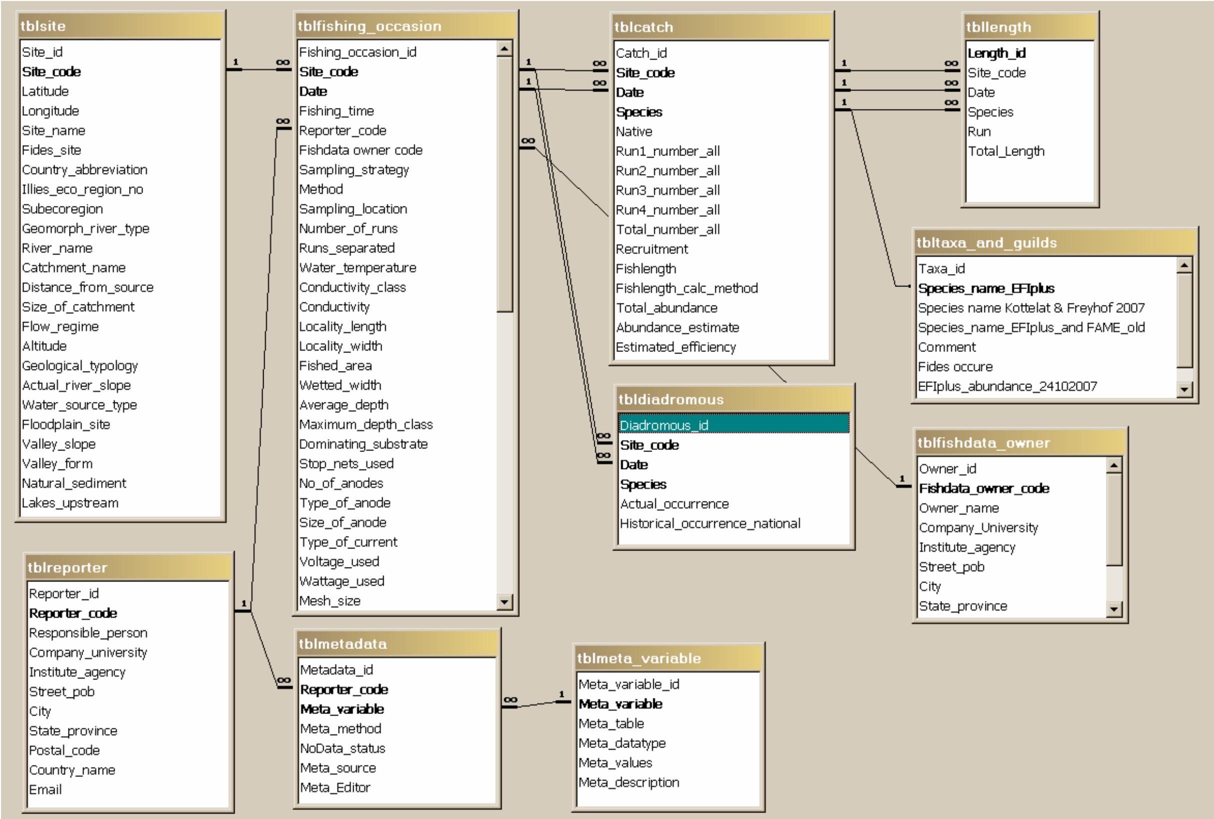

Figure 7: Structure of the EFI+ access database

The figure 8 shows the flow of data in EFI+ database

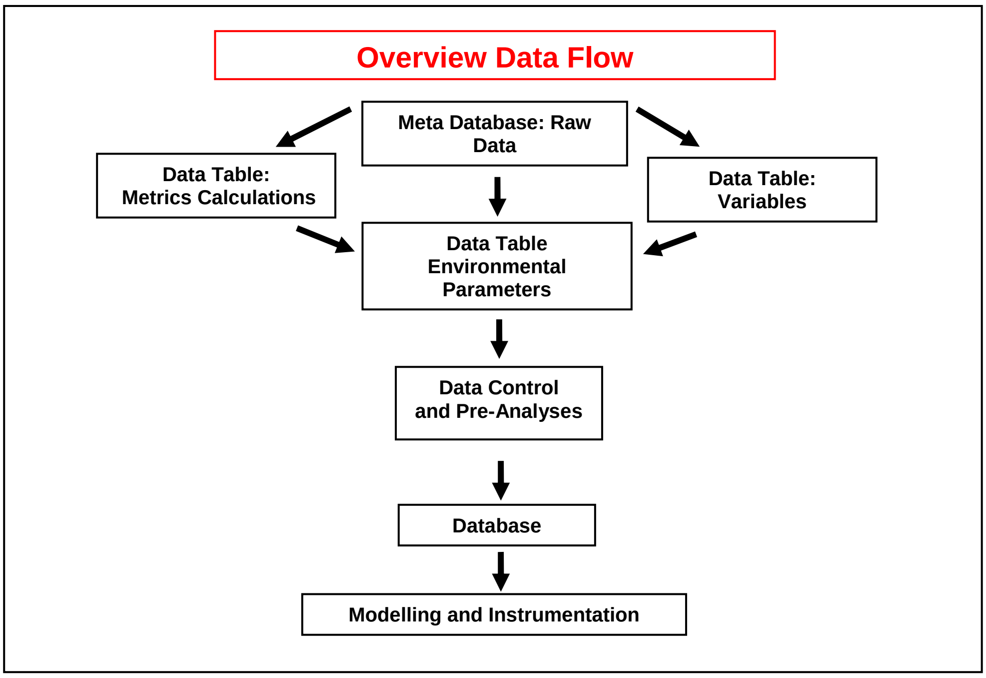

Figure 8: Flow of data of the EFI+ database

4. Description of Compilation of Austrian Data of the EFI+ Project

In this chapter the data of the Austrian dataset for the EFI+ project are described.

A general description of the European variables was not done because a complete variable description for all attending European countries was not available at the time when the thesis was written.

4.1. General Description of Variables

The variables can be divided into three different groups:

1. Environmental variables,
2. Pressure variables
3. Fish data (not used in this thesis).

The second group can be divided into four different pressure types: like hydrological pressures, morphological pressures, pressures on water quality, connectivity pressures.

Environmental and pressure data have to be collected by all European partners for each sampling site. These variables should be as consistent as possible and computed in the same manner Europe-wide.

All other variables will be calculated by all partners depending on the local or national scale.

To predict reference metric values the existing EFI uses environmental variables. Therefore, one of the tasks is to improve the data quality of environmental variables.

To increase the coverage of the environmental characteristics and to improve the reference models of the different partners new environmental variables have been defined, in particular for Mediterranean, Central/Eastern and Large Floodplain Rivers. The Environmental variables will be derived on a literature review, on the experience of specialists. The variables will be derived either locally or GIS-based. These environmental variables are needed to model the reference fish fauna on every site.

As mentioned before the environmental variables should not change depending on any existing pressure. For some variables, like river slope and wetted width the actual values will be considered since it will not be possible to gain information about the situation previous to major pressures within the 6-month-phase of data collection.

Several articles can be found, how pressures may significantly affect fish, their habitat, the entire river, fauna and flora and finally the surrounding landscape of a river. Here only a few of these resources should be mentioned. For example: Pressures which affect fish including altered land use (Wang et al. 2000), toxic pollutants (Breitholtz et a. 2001), acidification (Galloway 2001), sediment load (Hlass et al. 1998), increased salinity in dry regions (Hart et al. 1991), degradation of instream (Hall et al. 1996) and riparian habitat (Zalewski et al. 2001), flow regime alterations and dams (Lessard & Hayes 2003), loss of longitudinal and lateral connectivity (Holcik 2003), and organic and nutrient load (Degerman et al. 2001). Finally biological pressures may also be important, as introduction of fish species (Garcia-Berthou & Moreno-Amich 2000), excessive weed growth and weed cutting (Garner et al. 1996), fishing and stocking (Cowx & Gerdeaux 2004).

Main Pressures on River Systems:

Zitek et al. (2006) provide the basis for the definition of relevant criteria which can be used to quantitatively model fish/pressure relationships.

After Zitek et al. (2006) the following pressures are thought to be most relevant for rivers:

- Impoundment
- Alteration of the natural flow regime (hydro peaking, water diversion)
- Reservoir flushing (is consitered as a critical short term impact altering water quality and natural morphological character)
- Land use (is consitered to be an important indicator being indirectly related to many kinds of impacts)
- Alteration of the natural morphological character (canalization)
- Alteration of water quality (pollution)
- Loss of lateral, longitudinal and vertical connectivity
- River bed degradation
- Fish eaters, stocking, fishing pressure and alien fish species are additionally consitered as potential factors that can influence the fish fauna at a given site.
- Multiple/cumulative impacts
- Shipping

In the following figure 9 the interrelations among different pressures are shown:

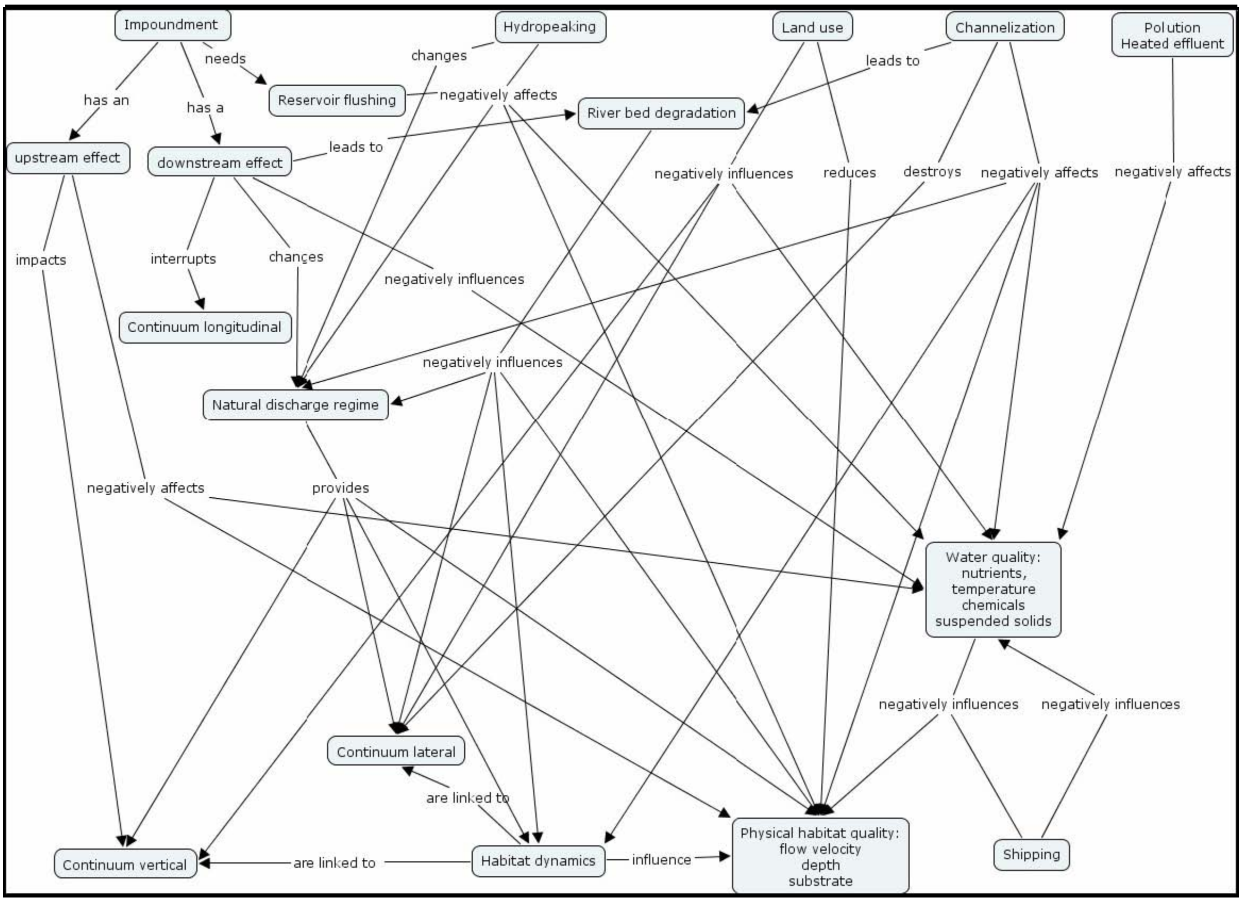

Figure 9: A theoretical scheme to organize pressures and effects relevant for Austrian rivers, (Zitek et al. 2006)

Jungwirth et. al. (2003) described in several examples the historical and up to date multidiversal anthropogenic influences and pressures on rivers and the impacts on natural river habitats and the fish fauna.

The main problems are: hydro peaking, interruption of the longitudinal continuum, interruption and prevention of dynamic processes, decrease of the structural diversity of the river habitats (structural heterogeneity), the separation of the unity of the river and flood plane, surrounding land connectivity (the lateral connectivity, back waters and dead arms), accumulation of fine sediments and sand, covering and colmation of bed sediments, alteration of the groundwater, alteration of the bed load, riverbed incision, flushing of impoundments and reservoirs, change of the temperature regime, water abstraction, and loss of free flowing river.

When these disturbances meet other anthropogenic disturbances like, eutrophication and many others it can lead to the development of extreme biotopes. In these biotopes only specialized and opportunistic species can survive.

In this chapter all variables for the specific data collection (data sources) are described. The following variables are the content of the main EFI+ Acces Tables: Catch Length, Fish Occasion. and Site.

4.1.1. Environmental Variables in Detail

Altitude

Description:

The variable altitude describes on how many meters above sea level of the sampling site is located. Unit: [m]

Data source:

The variable altitude is maintained from two different Austrian sources. The Austrian map ("Österreichische Karte 1:50.000 ÖK 50") and over a digital terrain elevation model (Digitales Geländehöhenmodell DGM).

Lakes Upstream

Description:

The variable describes if natural lakes present above the site. The answer can be Yes or No. It is only applicable if the lake affects the fish fauna of the site, e.g. by altering thermal regime or flow regime. The water WDF definition of lakes of more than 50 ha is used. If there are artificial lakes (as e.g. fish ponds upstream) these are pressures and must not be consitered in environmental variables.

Data source:

This information is provided by the "Umweltbundesamt Austria" (shapEFIle). It is the expert authority of the federal government in Austria for environmental protection and environmental control.

A helping variable is used over a "actual asset method" and is the base for the expert analyses.

EFI+ Criteria	**Lakes upstream**	Are there natural lakes present upstream of the site? Answer Yes or No. Only applicable if the lake affects the fish fauna of the site, e.g. by altering thermal regime, flow regime or providing seston. Use water frame directive definition of lake: more than 50ha .	yes	no
Helping Variable	SEE	IST- Bestandsanalyse - Waterbody ist ein See - Basis für Experteneinstufung	1	0

Distance from Source

Description:

Distance from source in kilometres to the sampling site measured along the river. In the case of multiple sources, the measurement shall be made to the most distant upstream source.

Data source:

Austrian maps (ÖK 50) or over a „GIS Tool" in combination with the the „Berichtsgewässernetz des Bundes".

Natural Flow Regime

Description:

Describes the type of the natural flow regime.

Where: Permanent: never drying out. Summer dry: drying out during summer. Winter dry: drying out during winter (e.g. some alpine or nordic rivers) Intermittent: Can be dry in any period of the year, otherwise it's summer- or winter dry. In Austria only the Permanent flow regime exists

Data source:

IHG database. BOKU - UNIVERSITY of NATURAL RESOURCES and APPLIED LIFE SCIENCES Institute of HYDROBIOLOGY and AQUATIC ECOSYSTEM MANAGEMENT Department of WATER, ATMOSPHERE and ENVIRONMENT

Geology

Description:

Describes the main dominating geological category. (e.g. Siliceous, calcareous or organic

Data source:

IHG database or geological maps.

Water Source Type

Description:

This variable is based on hydrograph of the river next to the sampling site. Four different types are use: Glacial, nival, pluvial or groundwater (all of them must be dominant).

Data source:

Mostly from the Hydrological Atlas in combination with tree different helping variables. Most important questions: in which month is the middle maximum discharge, appearance of year highest flooding and appearance of highest discharge?

Helping Variables:

REGIME_AKT_ANF: classification of the actual discharge regime: glazial, glazio-nival, moderate nival, nivaler Übergang, Nivoglazial, nivo pluvial, pluvial, pluvial nival A, pluvio, nival B Abflussmax_Monat_HAO: Seasonality of the monthly discharge, July - strong, June-strong, June – middle, May – middle, winter – weak, February – middle, March – middle, Spring – weak, April – strong, April middle.

Mader_Wimmer_Steidl: GLA8 – Glazial with discharge maximum in August, GLA7 Glazial with discharge maximum in July, NIG8: Nivio Glazial with discharge maximum in August, NIG7: Nivio Glazial with discharge maximum in July, NIV:6: Nivale Regime in June.

EFI+ Criteria	**Water source type**	based on hydrograph of the river next to the sampling site.	glacial	nival	pluvial
Helping Variable	REGIME_AKT_ANF	Einstufung der aktuellen Abflussregime	glazial, glazio-nival,	gemäßigt-nival, nival, nivaler Übergang, nivo-glazial, nivo-pluvial	pluvial, pluvio-nival A, pluvio-nival B,
Helping Variable	Abflussmax_Monat_HAO	Saisonalität der Monatsabflüsse - regionalisiert.	Juli stark	Juni stark, Juni mittel, Mai stark, Mai mittel,	Winter schwach, Februar mittel, März mittel, Frühling schwach, April stark April mittel
Helping Variable	Mader_Wimmer_Steidl	GLA 8 - Glazial mit Abflussmaximum im August, GLA 7 -Galzial mit Abflussmaximum im Juli, NIG 8 - NIvo Glazial mit Aflussmaximum im August, NIG 7 - NIvo Glazial mit Aflussmaximum im Juli, NIV 6 - Nivales Regime Juni	Gla 8, Gla 7	NIG 7, NIG 6, NIV 6 , GEN 6, GEN 5, NUE 5, NUE 4, WIN, HNI, NIP	PUE 4, PUE 3, SOP, PLN, WIP

Actual River Slope

Description:

Given as slope of stream bed along stream expressed as per mill, m/km (‰). The slope is the drop of altitude divided by stream segment length. The stream segment should be as close as possible to 1 km for small streams, 5 km for intermediate streams and 10 km for large streams.

Data source:

Calculated in GIS over maps with scale of preferably 1:25 000.

Valley Slope

Description:

Length of the valley between two contour lines.

Data source:

Calculated in GIS over maps.

Size of Catchment

Description:

Discribes the size of the catchment (watershed) upstream of the sampling site (km^2). Classes; <10, <100, <1000, <10000, >10000 km^2. (i.e. 0-9, 10-99, 100-999, 1000-9999,10000-).

Helping Variables:

Segment Length: Length of a segment specified after the catchments area size.

(1000, 5000, 10000

Catch _class: catchments area classes: 1 km for small rivers (catchment <100 km^2),

5 km for medium-sized rivers (100-1000 km^2) and 10 km for large rivers (>1000 km^2)

EFI+ Criteria	**Size of catchment**	Size of the catchment (watershed) upstream of the sampling site (km^2).			
Helping Variable	Segment length	Länge des Segments nach EZG Groesse	1000	5000	10000
Helping Variable	Catch_class	Einzugsgebietsklasse	100	100-1000	>1000

Catchment Name

Description:

Indicates the main catchment (based on ICES map), where river discharges into sea. Three catchment areas are known in Austria: Danube, Rhine and Elbe.

Data source:

As data source the Austrian hydrological Atlas and GIS is used.

Floodplain

Description:

Presence of a former floodplain: Answer is yes or no (e.g. significant area of adjacent landscape flooded at least every 10 years).

Data source:

Old maps, reports and expert judgment, consitered basically as a pre - categorization and then a process of elimination over the valley form and the valley slope is used. Finally sites from the ANF project were used.

Helping Variables:

ANF: Categories for pre-categorization. A; B means Yes and C, P means No

EFI+ Criteria	**Floodplain**	Presence of a **former floodplain:** yes, no (e.g. significant area of adjacent landscape flooded at least every 10 years), (data source: old maps, reports, expert judgement)	yes	no
Helping Variable	ANF	Kategorien	A, B	C, P

Valley Form

Description:

Four different Valley forms are defined (1): V-shape, (2) gorges, (3) U-shape, (4) plains.

Data source:

For deriving the valley slope for the Austrian Sampling points, Crane's proposed methods was partially followed. The river segments were extracted, their endpoints sea level height from DEM (10m grid) were queried and derived a sinuousity index from riversegment and euclidean distance ratio. Then the dataset was grouped the in classes of segment lengths and sinuousity and some samples were checked by interpretation of topographical maps. Finally the riversegments were smoothed with tolerances of 500, 625, 1250, 2500 and 5000 to represent valley axis. Best fitting axis were selected by classification in the previous step and in the end manually checked for errors. The valley width was not model dependent from catchment size because there would have been to many exeptions.

Scheme of ANF, then the valley form is calculated over the valley slope and the geographical river type on MIRR- Points and expert judgement.

Helping Variables:

Valley form_ANF: gorge: Klamm Schlucht, V-shaped valley: Kerbtal Sohlenkerbtal, U-shaped valley: Trogtal, Muldental (wide open valley with sites flaring out), Sohlental, plain: Talebene

EFI+ Criteria	**Valley form**	(1): (2) (3) , (4)	gorges,	V-shape,	U-shape	plain
Helping Variable	Valley form_ANF		Klamm/ Schlucht	Kerbtal, Sohlenkerbtal	Trogtal, Muldental, Sohlental	Talebene
Classification	MIRR	Erhebungsbögen Datenbank	Klamm/ Schlucht	Kerbtal, Sohlenkerbtal	Trogtal, Muldental, Sohlental	Talebene

Classification: MIRR: data entry form from databank.

Geomorphological River Type

Description:

This information is provided in four categories: naturally constraint without mobility (riverbed is fixed), braided, sinous and two meandering classes (regular and tortous, different in sinousity index). Situation before any major human control of river bed.

Data source:

Sinuosity Index river length and valley axis and several helping variables See valley form.

Helping Variables:

CHA_AM_POT: ANF categorization, morphological historic-potential river type: naturally constraint; gestreckt ,Talmeander, braided: Furkation, gewunden, sinous: pendelnd, regular meandering: Mäander, tortous meandering: Mäander.

MIRR: hist rtype in classes: naturally constraint; 1-gestreckt , 8-Talmeander, braided: 2-Furkation, 5-gewunden, sinous: 3-pendelnd, regular meandering: 7-Mäander, tortous meandering: 7-Mäander

Sinuosity Index: Classification: naturally constraint; 1.0, braided: 1.2, sinous: 1.5, regular meandering: 1.5, tortous meandering: 2.0.

EFI+ Criteria	**Geomorphological river type**	4 categories to be selected: naturally constraint without mobility (riverbed is fixed), braided, sinous,and 2 meandering classes (regular and tortous, different in sinousity index).Situation before any major human control of river bed!	naturally constraint	braided	sinous	regular meandering	tortous meandering
Helping Variable	CHA_AM_POT	ANF Kategorisierung	gestreckt Talmäander	Furkation , gewunden	pendelnd	Mäander	Mäander
Helping Variable	MIRR	hist_rtype in Klassen	1 gestreckt 8 Talmäander	2 Furkation , 5 gewunden	3 pendelnd	7 Mäander	7 Mäander
Helping Variable	River length on segment						
Helping Variable	Segment length						
Helping Variable	Sinousity index:	Klasseneinteilung wie?	1	1.2	1.5	1.6	2

Naturally Dominant Sediment

Description:

This variable is divided in five categories: organic, silt, sand, gravel-pebble-cobble, boulder-rock; Situation before major changes of sediment conditions, always for the dominating substrate. At large rivers, consiter dominant sediment in the potamic zone with weak to medium water depths.

Data source:

Expert judgment.

4.1.2. Variables Describing the Sampling Methods

Sampling Strategy

Description:

Definition of how the section was sampled. The whole river width (whole) or only parts of the river (partial).

Method

Description:

Defines if electric fishing was carried out by wading, boat or mixed (sites sampled with both - wading and boat).

Fished Area

Description:

Area of the section that has been definitely sampled (sampled length * sampled width) given in m^2.

Floodplain

Description:

Defines if the sample was taken from main channel, floodplain or mixed.

4.1.3. Variables Describing the Location, Name of Site and Date of Fishing

Site code XX_Sitecode

Description:

This is a unique reference number per sampling site. (International country code (XX, e.g. AT for Austria) plus national code separated by "_")

Date

Description:

The sampling date is provided in the format: Day/Month/Year e.g. 08/08/1974.

Latitude

Description:

The Latitude is given in WGS 84 decimal format, 6 digits behind the comma.

Data source:

GPS and digital maps

Longitude

Description:

The Logitude is given in WGS 84 decimal format, 6 digits behind the comma.

Data source:

GPS and digital maps

Latitude_CCM

Description:

This is a Modified Latitude. Derived by GIS from the CCM river network

Data source:

Derived by GIS from the CCM river network

Longitude_CCM

Description:

This is a Modified Longitude. Derived by GIS from the CCM river network

Data source:

Derived by GIS from the CCM river network

River Name

Description:

This is the official name of the river used in the country. In the case of a trans-boundary river, the name of the river is taken from the country where the river has its mouth. (either to the sea or to the next river downstream).

Site Name

Description:

This is a location name and it is indicating a nearby town or village.

Fides

Description:

Describes if the Site exists already in FIDES. Answer can be yes or no

4.1.4. Hydrological Pressures Variables in Detail

<u>Impoundment</u>

Description:

Describes if the natural flow velocity is reduction on site due to impoundment.

Answer: No, Weak or Strong

Data source:

1. Variables: Select impoundment_MIRR, ANF-Impoundment, IST-Bestand-STAU,
2. Select impoundment = 1, then the Impoundment is strong
3. If "Stau" is mentioned in the text: Oberwasser strong, Unterwasser weak
4. ANF – continuous (durchgehend) + Risk (Risiko) 3 = strong
5. ANF – continuous (durchgehend) + Risk (Risiko) 2 = check
6. ANF – partly in the area (Bereichsweise) + Risk (Risiko) 3 = weak
7. Elswise: Expert judgement

Helping Variables:

EFI+ Criteria	**Impoundment**	Natural flow velocity reduction on site due to impoundment	no (no impoundment)	weak	strong
Help. Var.	Select_ impoundment	Staupunkt/ ja, nein	0		1
Help. Var.	Impoundment _ ANF		nicht vorhanden	bereichsweise, lokal	durchgehend bis fast durchgehend
Help. Var.	STAU	Risiko IST-Bestandsanalyse, 1 bis 3, Kat. 2 - Risiko nicht einschätzbar	1	2	3

<u>Hydropeaking</u>

Description:

This variable indicates if the site is affected by hydropeaking. Anwer: Yes or No

Data source:

Variablen: ANF-Hydropeaking, IST-Bestand-SCHWALL,

1. ANF – durchgehend = yes
2. ANF – bereichsweise = yes

3. ANF – lokal – check nötig
4. Risiko IST Bestand 3 = yes
5. Risiko IST Bestand 2 = yes
6. Risiko IST Bestand 1 und ANF aber ja = check
7. Sonst Expert judgement

Helping Variables:

EFI+ Criteria	**Hydropeaking**	Site affected by hydropeaking	no (no hydropeaking)	yes		
Helping Variable	Hydropeaking_ANF		nicht vorhanden	bereichsweise	lokal	durchgehend bis fast durchgehend
Helping Variable	SCHWALL	Risiko IST-Bestandsanalyse, 1 bis 3, Kat. 2 - Risiko nicht einschätzbar	1	2	3	

<u>Water Abstraction</u>

Description:

Gives answer to the question if the site affected by water flow alteration/minimum flow (water abstraction)

Helping Variables:

EFI+ Criteria	**Water abstraction**	Is the site affected by water flow alteration/minimum flow (water abstraction)	no	weak to medium (less than half of the mean annual flow)	strong (more than half of mean annual flow)
Helping Variable	Residual_length	MIRR RW Info		bei MQ perc über 50	bei MQ perc unter 50
Helping Variable	Residual_flow_ANF		nicht vorhanden	bereichsweise, lokal	durchgehend bis fast durchgehend
Helping Variable	RESTWASSER	Risiko IST-Bestandsanalyse, 1 bis 3, Kat. 2 - Risiko nicht einschätzbar	1	2	3

<u>Water Use</u>

Description:

Is the driving force of pressure: hydro power, irrigation, drinking water, snow production, fish ponds, cooling for thermal/nuclear power plants etc.) Answer No orr: HP (hydropower), I (Irrigation), DW (Drinking Water), SP (Snowproduction), FP (Fishponds), CW (Cooling water); IW (Industiral water), OT (others).If Yes or weak strong with at the variables Impoundment, Hydro peaking, or Water abstraction then HP (hydropower).

Data source:

If Schwall Restwasser oder Stau then the variable hydropower is used. Others are in Austria not existent

Hydrograph Modification

Description:

Seasonal hydrograph modification due to hydrological alteration (e.g. water storage for irrigation, hydropower,…) Answer: Yes or No

Data source: Expert Judgement

Helping Variables:

EFI+ Criteria	**Hydrograph modification**	Seasonal hydrograph modification due to hydrological alteration (e.g. water storage for irrigation, hydropower,…)	no	yes
Helping Variable	ANF	REGIME_POT		
Helping Variable	ANF	REGIME_AKT		
Helping Variable	HYDROLOGIE	Risiko IST-Bestandsanalyse, 1 bis 3, Kat. 2 - Risiko nicht einschätzbar	1	2, 3
Helping Variable	ANF	Residual_flow	nicht vorhanden	bereichsweise, lokal, durchgehend bis fast durchgehend

Temperature

Description:

Is there an impact on water temperature. Answer Yes or No

Data source: Expert Judgement

Flow Velocity Increase

Description:

Is there an impact on flow conditions (mean velocity) due to channelisation, floodprotection, etc.

Data source:

Variables: RT-Index_MIRR,

1. River type index 5 und kein Stau = yes

2. River type index 1 oder Stau = no

3. River type index 3 und durchgehend. Longbank = yes

4. Additional Morphology – channelisation: strong = yes, intermediate only if embankment continous = 5.

6. Additional Extpert judgement

Helping Variables:

EFI+ Criteria	**Flow velocity increase**	Is there an impact on flow conditions (mean velocity) due to channelisation, floodprotection, etc.	no	yes
Helping Variable	ANF	Longbank_protection		bereichsweise, lokal, durchgehend bis fast durchgehend
Helping Variable	MIRR	Cha_rtype - Veränderung des Flusstyps - Index	1	5
Helping Variable	MIRR	bw_sohle Bewertung der Sohle nach NÖMORPH		Klasse 4: durchgehend Pflasterung verfugt oder Beton
Classification				100 an Flüssen wie Fischa, Piesting, Ybbs, Erlauf etc.

Reservoir Flushing

Description:

Is the fish fauna affected by flushing of reservoirs upstream of the site? Answer Yes or No

Data source: Expert Judgement

Sedimentation

Description:

Input of fine sediment (mainly mineral input; bank erosion, erosion from agricultural land, etc. Answer: No, weak (slight reduction of sediment porosity), Medium, high (coarse sediment clogged)

Data source:

Origin data: Shape File, Bodenabtrag des HAO;

Method: Puffer 10 km long und lateral 1 km; weighted mean per Puffer

1/no: Vegetationsarme Flächen; kein Bodenabtrag; bebaute Siedlungsflächen

2/weak: Wald; sehr geringer Bodenabtrag

3/medium: mittlerer Bodenabtrag

4/high: hoher Bodenabtrag, sehr hoher Bodenabtrag

Final classification of weighted mean per Puffer:

≤1.5: no

≤ 2: weak: slight reduction of sediment porosity

≤ 2.5: medium:

2.5 high: coarse sediment clogged

Helping Variables:

EFI+ Criteria	**Sedimentation**	Input of fine sediment (mainly mineral input; bank erosion, erosion from agricultural land, etc.)	no	weak (slight reduction of sediment porosity)	medium	high (coarse sediment clogged)
Helping Variable	Erosionsindex HAO	Karte Bodenabtrag durch Wasser	Gletscher (kein Bodenabtrag), Vegetationsarme Flächen (kein Bodenabtrag), Bebaute Siedlungsflächen (kein Bodenabtrag) kein Bodenabtrag	Wald (sehr geringer Bodenabtrag, 150kg/ha/a), sehr geringer Bodenabtrag (0-1000 kg/ha/a)	mittlerer Bodenabtrag (1000 - 5000kg/ha/a)	hoher Bodenabtrag (5.000 - 10.000 kg/ha/a), sehr hoher Bodenabtrag (>10000kg/ha/a)

4.1.5. Morphological Pressures Variables in Detail

Channelisation

Description:

This variable describes an alteration (change) of natural morphological channel plan form (intensity of straightening)

Data source:

ShapEFIles from GIS, straightened hoarmonisation of the FF-Index Variables: River type index for MIRR and ANF, Morphological_condition (optional raramter: old classification from FIDES) for not MIRR and not ANF-Daten

1. River type index 5 = straightened
2. River type index 3 = intermediate
3. River type index 1 = no
4. Morphological_condition FIDES 4 oder 5 = straightened
5. Morphological_condition FIDES 3 oder 5 = intermediate
6. Morphological_condition FIDES 1 oder 2 = no
7. Additional Expert judgement

Helping Variables:

EFI+ Criteria	**Channelisation**	Alteration (change) of natural morphological channel plan form; intensity of straightening	no	intermediate	straightened
Helping Variable	MIRR	Cha_rtype - Veränderung des Flusstyps - Index	1	3	5
Helping Variable	FAME	Morphological Condition site	most of nat. channel maintained, all habitat types present (2), negligible morphological alteration (1)	channelised, some natural habitats missing (3)	canal (5), channelised most nat. habitats missing (4),

Channelisation Cross Section

Description:

Describes an alteration of the cross section.

Data source:

Quarry trough IHG and GIS databank

Variables: River Type index for MIRR und ANF, ANF Kategorisation,

1. River Type 5 = technical

2. ANF all A = no alteration

3. B, P parts with RT 1 = no alteration

4. ANF B, P and C-Strecken die RT 3 = intermediate

5. Expert judgement

Helping Variables:

EFI+ Criteria	**Channelisation**	Alteration of cross section	no	intermediate	technical crossec./U-profil
Helping Variable	MIRR, ANF	River Type Index	1	3	5
	ANF	Kategorien ANF	A, B, P	B, P, C	

Channelisation - Instream Habitat

Description:

Describes an alteration of instream habitat conditions.

Data source:

Quarry trough IHG and GIS databank. Variables: Noemorph tooting for MIRR Data, ANF – categorisation, Morphological_condition (optional Parameter – old categorisation from FIDES) but not MIRR and not ANF-Data

1. Noemorph-MIRR tooting – classification

2. ANF – A Strecken – no alteration

3. ANF C – Sections with throughout holdup (Stau) – high

4. ANF C – Sections with throughout Longbank = high

5. ANF B – Sections with local Long-bank – no alteration

6. others ANF B – Sections with = intermediate

7. ANF P expert judgement

8. ANF C with sectored holdup (Stau), sectored Longbank – expert judgement –

9. Morphological_condition FIDES 4 oder 5 = high

Morphological_condition FIDES 3 = intermediate

Morphological_condition FIDES 1 oder 2 = no

10. Expert judgement

Helping Variables:

EFI+ Criteria	**Channelisation**	Alteration of instream habitat conditions	no	intermediate	high
Helping Variable	MIRR-Nömorph	Verzahnung	1 - 1,75	2 - 2,75	3 bis 4
Helping Variable	ANF	Kategorien ANF	A, B	B	C
Helping Variable	ANF	Impoundment			durchgehend
Helping Variable	ANF	Longbank_protection	local		durchgehend
Helping Variable	FAME	Morphological Condition site	most of nat. channel maintained, all habitat types present (2), negligible morphological alteration (1)	channelised, some natural habitats missing (3)	canal (5), channelised most nat. habitats missing (4),
Helping Variable		Morphologiekart. Salzburg, Vbg, Stmk,			

Cannelisation - Riparian Vegetation Close to Shoreline

Description:

Describes an alteration of riparian vegetation close to shoreline.

Data source:

Quarry trough IHG and GIS databank. Variable: Percentage of forest in local buffer at site (75m up- and 75m downstream of site, lateral 50m left and right) Data: SINUS landuse classification

Helping Variables:

EFI+ Criteria	**Channelisation**	Alteration of riparian vegetation close to shoreline	no	slight	intermediate	high (no riparian vegetation)
Helping Variable	SINUS-Datensatz - Landuse Kategorien	loc_con_sin1	75-100% remaining	50- 75% remaining	25 - 50% remaining	<25% remaining

Cannelisation – Artificial Embankment

Description:

Describes the force of an artificial Embankment

Data source:

Quarry trough IHG and GIS databank. Variables: Noemorph-Bank for MIRR Data, Longbank protection for ANF passage.

1. Noemorph – Bank 1 bis 1,25 = no
2. Noemorph – Bank 1,5 bis 1,75 = slight
3. Noemorph – Bank 2 bis 2,75 = intermediate
4. Noemorph – Bank 3 bis 4 = high
5. ANF continous = high
6. ANF sectored = intermediate
7. ANF local = slight
8. ANF not existent = no
9. Expert judgement

Helping Variables:

EFI+ Criteria	**Channelisation**	Artificial embankment	no (natural shoreline)	slight (local presence of artifical material for embankment)	intermediate (continuous embankment but permeable (e.g. rip rap))	high (continuous, no permeability (e.g. concrete walls))
Helping Variable	ANF	Longbank_protection	nicht vorhanden	lokal	bereichsweise	durchgehend bis fast durchgehend
Helping Variable	M-Bank	MIRR	1 bis 1,25	1,5 bis 1,75	2 bis 2,75	3 und 4
Helping Variable	Ufer_links	Morphologiekart. Salzburg	ungesichert	punktuell gesichert	streckenweise	durchgehend gesichert

Floodprotection - Dykes For Flood Protection

Description:

This variable is a criterion for the following variable Floodplain and describes if a dyke for flood protection exist.

Data source: Quarry trough IHG and GIS databank verified with expert judgement.

1. ANF A and B = no

2. ANF C and P = yes

3. If Floodplain FIDES 1 = check

4. If Floodplain FIDES 5 = yes

Helping Variables:

EFI+ Criteria	**Floodprotection**	presence of dykes for flood protection	no	yes
Helping Variable	ANF	Kategorien ANF	A, B	C, P

If in the area A and B then its is a floodplain. (see 3.8.6.6)

Floodprotection - Connected Floodplain

Description:

Describes If the river has a former floodplain. The Proportion of connected floodplain are still remaining. (Floodplain = area connected during the flood).

Data source:

Quarry trough IHG and GIS databank verified with expert judgement. ANF – change of the class.

Helping Variables:

EFI+ Criteria	**Floodplain**	If the river has former floodplain - Proportion of connected floodplain still remaining.	no (= no former floodplain available)	> 50%	10-50%	less than 10 %	some waterbodies remaining
Helping Variable	ANF	Akt. Auwaldgroesse	Akt.Auwald-sehr gross	Akt.Auwald-gross	Akt.Auwald-mittel	Akt.Auwald-klein	Akt.Ausaum
Helping Variable	ANF	Pot. Auwald	Pot.Auwald-sehr gross	Pot.Auwald-gross	Pot.Auwald-mittel	Pot.Auwald-klein	Pot.Ausaum
Helping Variable	ANF	Change		bleibt gleich	Änderung um 1 Klasse	Änderung um 2 Klassen	Änderung um mehr als 2 Klassen
Helping Variable	Flussaue	Morphologiekart. Salzburg, Vbg, Stmk,					
Helping Variable	SINUS-Datensatz - Landuse Kategorie Wald	Für ganz Österreich					

4.1.6. Water Quality Pressures Variables in Detail

Waterquality - Toxic Substances

Description:

Describes if toxic priority substances are present (organic and nutrient appearance).

Data source:

Quarry trough DB and Rauchbuchl Excel file verified with expert judgement and shape file.

1. Risc 3 = Strong
2. Risc 1 = No
3. Risc 2 = Weak

Helping Variables:

EFI+ Criteria	**Water quality**	Toxic prioriy substances (organic and nutrient appearance)	no or very minor (e.g. atmospheric input far away, no contamination in the segment itself)	weak (important risk, link to particular substance)	high concentration (a clearly kown input)
Helping Variable	IST-Bestand	RISIKO_L1PSSO	1	2	3

Waterquality - Acidification

As there is no acidification in Austria therefore this pressure is not existent.

Waterquality - National Water Quality Index

Description:

This pressure National water quality index is divided into in 5 classes.(indication of classes (1 to 5; 5 = worst)).

Data source: Waterquality, Gewässergüte 2002, IST-Bestandsanalyse UBA, ShapEFIle

Helping Variables:

EFI+ Criteria	**Water quality**	National water quality index - in 5 classes, indication of classes (1 to 5; 5 = worst)	1	2	3	4	5
Helping Variable	UBA	Gewässergüte	I; I-II	II, II-III (potamal)	II-III (rhitral), III	III-IV	IV

Waterquality - Eutrophication

Description:

This Variable answers the question if there is any artificial eutrophication.

Variablen: Orthophosphat WGEV- measurement stations blended with Water bodies

Helping Variables:

Quarry trough DB data file verified with expert judgement and shape file.

Helping Variables:

EFI+ Criteria	**Water quality**	National water quality index - name	name of index			
EFI+ Criteria	**Water quality**	Is there an artificial eutrophication?	no	low	intermediate (occurance of green algae)	extreme (oxygen depletion, increase of primary production)
Helping Variable	Orthophosphat	0,001667 bis 0,215833	bis 0,04	0,04 - 0,075	ab 0,075	

Waterquality - Organic Pollution

Description:

Answers the question: is there is any organic pollution observed?.

Helping Variables:

Gewässergüte 2002, IST-Bestandsanalyse UBA, ShapEFIle

Helping Variables:

EFI+ Criteria	**Water quality**	Is organic pollution observed?	no	weak	strong
Helping Variable	UBA	Gewässergüte	I; I-II;	II; II-III (potamal),	II-III (rhithral), III,

Note: in Austria only the classes 1-3 exist.

Waterquality - Organic Siltation

No pressure in Austria.

4.1.7. Connectivity Pressures in Detail

Barriers catchment down:

Description:

Presence of downstream barriers.

Helping Variables:

EFI+ Criteria	**Migration barriers - catchment scale**	Presence of downstream barriers on the catchment scale	no	partial: migration possible for good swimmers (e.g. salmon) or for particular situations/years	yes: definite barriers for most species most of the time
Helping Variable	IST-Bestand, Querbauwerke				yes: für alle sites in Österreich

Barriers segment up:

Description: Presence of Barriers on segment level upstream

Helping Variables:

EFI+ Criteria	**Migration barriers - river segment scale**	Barriers on segment level upstream	no	partial	yes

Barriers segment down:

Description: Presence of Barriers on segment level downstream

Helping Variables:

EFI+ Criteria	**Migration barriers - river segment scale**	Barriers on segment level downstream	no	partial	yes

Barriers_number_river_segment_up:

Description: Number of Barriers on segment level upstream

Helping Variables:

EFI+ Criteria	**Migration barriers - river segment scale**	Number of barriers upstream	no = 0	yes = nr. of barriers

Barriers_number_river_segment_down:

Description: Number of Barriers on segment level downstream

Helping Variables:

EFI+ Criteria	**Migration barriers - river segment scale**	Number of barriers downstream	no = 0	yes = nr. of barriers	
Auswertung	wie oben	UBA und MIRR querbauwerkslayer			GIS-Auswertung für up- und downstream über Segment

Distance Barriers downstream

Description: Distance to next barrier in the segment – downstream

Helping Variables:

EFI+ Criteria	**Migration barriers - river segment scale**	Distance to next barrier in the segment - downstream	0 = no barrier	in km

Distance Barriers upstream:

Description: Distance to next barrier in the segment - upstream

Helping Variables:

EFI+ Criteria	**Migration barriers - river segment scale**	Distance to next barrier in the segment - upstream	0 = no barrier	in km	
Auswertung	wie oben	UBA und MIRR querbauwerkslayer			GIS-Auswertung Network Analyst (cutoff value= segmentlength/2)
Helping Variable	MIRR	d_next_b_up - Abstand von der nächsten Kontinuumsunterbrechung flussauf	nicht verwendet		d_next_b_up oder d_next_b_up2 verwenden? Update vom 040407
Helping Variable	MIRR	d_next_b_down - Abstand von der nächsten Kontinuumsunterbrechung flussab	nicht verwendet		d_next_b_up oder d_next_b_up2 verwenden? Update vom 040408

Data for All Connectivity Pressures:

Layer barriers from MIRR-project for MIRR-sites;Layer barriers from UBA

Data source: Quarry over about the amount of Barriers (30m Puffer an das Segment (Hawth's Tools – Count Points in Polygon) (Berichtsgewässernetz))

Quarry of the distance to the next barrier river (up/downstram) with Network Analyst (ArcGIS9.2) mit CutoffLength = Segmentlength upstream bzw. downstream

4.1.8. Data Acquisition for the EFI+ Central Database at European Scale

As for the EFI+ project a database of existing electro fishing data in European streams together with parameters and variables describing the sampling site is necessary it was important and to acquire many environmental descriptors and pressure variables for the sampling sites.

Environmental variables are used for modelling of references of the fish fauna on site level and pressure data allows identification of reference sites as well as developing pressure intensity indices.

The data acquisition was either done individually by the project partners themselves or additionally within a sub working task using a geographical information system (GIS) and European wide spatial databases for querying variables for all sites at once.

All EFI+ sites are linked to a pan-European river and catchment database (CCM – catchment characterisation and modelling, version 2 released in June 2007) which was developed at the Joint Research Center in Ispra, Italy. Many catchment relevant variables are yielded within this data but by using the hydrologic coded structure of the database other pressure and environmental variables could be calculated and be aggregated for individual catchments of EFI+ sites.

Other data was acquired from the European Soil Database (ESDB, Quarry CORINE land cover (CLC2000, European Environmental Agency), pan-European Forest Map 2000 (JRC, Ispra), road layer of Euro Global Map (provided by EUROSTAT), and worldclim database (worldclim.org).

5. Methods

5.1. General Description of Data Mining

Our human society produces in the different areas of the economic and scientific world day after day huge amounts of raw data. As for this data most of them hide potential valuable and vulnerable information for future economic and scientific decisions and planning. The main problem is that this raw data are useless without any information extraction treatment. Therefore, the science provides different useful technologies. One of these techniques is data mining.

Data mining can be a useful and powerful instrument with great possibilities and the potential to help focus on the most important information in collected data. Data mining involves sorting and analyzing large amounts of data and picking out relevant information. Many companies, intelligence organizations, and financial analysts usually use it. Finally, it is increasingly and widely used in the sciences to extract information from enormous data sets generated by modern experimental and observational methods. (Doug A., accessed in 07/2007, http://www.eco.utexas.edu/~norman/BUS.FOR/course.mat/Alex/, accessed in 07/2007)

Frawley et al. (1992) defined data mining as the nontrivial extraction of implicit, previously unknown, and potentially useful information from data.

Another expression for data mining was given by Hand et al. (2001). He expressed data mining it as the science of extracting useful information from large data sets or databases.

The third essential expression presents Data Mining as knowledge discovery. It is a computer-assisted process of digging through and analyzing enormous sets of data and finally the extracting of the possible meanings of the data. Data mining tools like “SPSS”, and others predict behaviours and future trends, allowing business and scientific institutions to make active and knowledge-driven decisions and furthermore future planning.

These tools can answer questions that traditionally were too time consuming to resolve. They scan databases for hidden patterns and find predictive information that experts may miss because it lies outside their expectations.

Many analysts separate data mining software into two groups. The first group, Data mining tools provide a number of techniques that can be applied to any statistical problem. The second group, data mining applications, uses techniques insite an application solve a specific statistic problem.

Data mining can be a useful but also dangerous information instrument. As many data provide many information's of human live this data can be also used ethic questions. In the following many data can be used for criminalisation and discrimination purposes.

5.2. Descriptive Statistics

Descriptive statistic is used to convey the important aspects of the distribution of large collected data sets. and present quantitative descriptions of the data in a manageable form.

(Davidson College, 2002; http://www.bio.davidson.edu/Courses/Bio111/Bio111LabMan/Lab%207.html accessed in 09/2007)

There are three major characteristics of a single variable that that can be described:

1. the central tendency and

2. the spread of your data

3. the dispersion

The main task of descriptive statistic is to provide simple summaries about datasets, samples and measures. Data can be summarized in the following way: graphical description (graphs) tabular description (tables) and summary statistics (calculated values)

To summarize or explain a quantity like a length, weight or age, it is common to answer the first question with the arithmetic mean, the median, or the mode.

In this thesis, descriptive statistics is used :

• to control data about their plausibility and missing values per analyse of frequencies, cross tables, tables and bar graphs,

• to compare data with explorative statistics, in tabular and graphically form.

Type of data:

Numerical data: Data measured or identified on a numerical scale, can be analysed using tables, charts, histograms and graphs.

Categorical data: Data, which is placed into categories such as age groups. (non numerical data) can be analyses by using regression analyses, analyses of variance et cetera.

5.2.1. Arithmetic Mean

A simple measure of the central tendency of the data is the mean (or average): mean = sum of all the data / sample size (often called n)

5.2.2. Median

A median is described as the number separating the higher half of a sample, a population, or a probability distribution, from the lower half. The median of a finite list of numbers can be found by arranging all the observations from lowest value to highest value and picking the middle one. If there is an even number of observations, the median is not unique, so one often takes the mean of the two middle values.

For example, with the data set (1, 2, 2, 5), n=4, the mean is (2+2) / 2=2

5.2.3. Range

The simplest measure of the spread of your data is the range. It describes the distance between the most extreme data values. The range does not describe how frequent these extreme values are present in the data set.

The formula for calculating the range is:

range = value of maximum data point minus value of minimum data point

For example, with the data set (1, 2, 2, 5), the range is 5 - 1 = 4.

5.2.4. Variance

The variance of your data is a measure of the spread of the data. It will take into account: both deviations of the data (away from the mean) and how frequently these deviations occur.

For each data point, the mean is subtracted from each data point, and then this value is squared. The squared values are added together and divided by either n or n minus 1. If the entire population was sampled, then it is divide by n. If the subset of a population sampled it is divide by n-1.

In the following example the entire population was sampled.

The formula for calculating variance is:

variance = the sum of (each data point minus the mean)^2 ÷ sample size

For example, with the data set (1, 2, 2, 5) $(1 - 4)^2 + (2 - 4)^2 + (2 - 4)^2 + (5 - 4)^2 = 18$

The variance is 18 / 4 = 4.5

5.2.5. Standard Deviation

The standard deviation is the square root of the variance. It reflects both the deviation from the mean and the frequency of this deviation.

The standard deviation is more often used then the variance because the scale of the variance tends to be larger than the scale of the raw data, while the standard deviation is on the same scale as most of the data.

The formula for standard deviation is:

Standard deviation = sq root (variance)

For example, with the data set (1, 2, 2, 5),

The standard deviation is the square root of 4.5 = 2.12

5.2.6. Standard Error of the Mean

The standard error of the mean describes the deviation from the mean and the frequency of this deviation. It also takes into account the size of the set.

The formula for standard error is:

standard error = sq root (variance / n) (n= sample size)

For example, with the data set (1, 2, 2, 5)

the standard error is the square root of 3 / 4 = 0.53.

For better understanding why standard error is a useful statistical description, another data is used set where the variance was 4.5 but n = 20.

Standard Error = square root of 4.5 / 20 = 0.1.

The same variance of 3 gave different standard errors (if n=4: 0.53 versus if n=20: 0.1) due to the difference in sample size. However, a look at standard error and standard deviation, shows that the standard error has taken the sample size twice into account.

In reality, scientists use the standard error to make their data look better than they are. The standard error is a statistical analysis of one set of data used as actually repeated the same experiment many times and gotten a range of means.

In other words, standard error is a statistical approach that attempts to look at the variance of these imaginary range of means and determine the variance of these means.

It expresses that it is > 95% sure, if the particular experiment is repeated another time, the mean value would fall within a certain range.

5.2.7. Bar Graph

A bar graph, also known as a bar chart, is a chart with rectangular bars of lengths usually proportional to the magnitudes or frequencies of what they represent.

A bar chart is normally used for comparing two or more values. The bars can be vertically or horizontally oriented. Also, different groups can be indicated using different coloured bars, or bars with different patterns. Also a stretched graphic can be used instead of a solid bar.

The bar chart is possibly the invention of the Scots engineer and economist William Playfair (1759-1823). A bar chart was used in his work *The Commercial and Political Atlas* (London, 1786) (Wikipedia, http://en.wikipedia.org/wiki/Bar_graph, accessed in 07/2007).

5.2.8. Histogram

A histogram is used in Statistics as a graphical display of tabulated frequencies.

It is the graphical version of a Table that shows what proportion of cases fall into each of several or many specified categories.

The histogram differs from a bar chart in the way that it is the area of the bar that denotes the value, not the height.

This is a crucial distinction when the categories are not of uniform width (Lancaster, 1974). The categories are usually specified as non-overlapping intervals of some variable. The categories (bars) must be adjacent.

The word histogram is derived from Greek word: histos 'anything set upright' (as the masts of a ship, the bar of a loom, or the vertical bars of a histogram); gramma 'drawing, record, writing'.

The histogram is one of the seven basic tools of quality control, which also include the:

● Pareto chart

● check sheet

● control chart

- cause-and-effect diagram
- flowchart
- and scatter diagram.

A generalization of the histogram is the kernel smoothing techniques. This will construct a very smooth probability density function from the supplied data. (Wikipedia, http://en.wikipedia.org/wiki/Histogram, accessed in 07/2007)

5.2.9. Error bar

An error bar chart is a graphic way of summarizing the mean scores for a group of data. It displays graphically the 95% confidence interval of the mean for groups of cases. The boxes in the middle of the error bar represent the mean score. The "whiskers" represent the 95% confidence interval.

It also allows to graphically illustrate actual errors, the statistical probability of errors, or a general approximation or "spread" in your data. Examples might include experimental errors in measurement or atypical data points in comparison to the rest of the data.

5.2.10. Box Plot

In descriptive statistics, a boxplot (also known as a box-and-whisker diagram or plot or candlestick chart) is a convenient way of graphically visualisation of the five-number summary. The five number Summary consists of the smallest observation, lower quartile, median, upper quartile, and largest observation; in addition, the box plot indicates which observations, if any, are consitered unusual, or outliers.

The box plot was invented in 1977 by the American statistician John Tukey. box plots are able to visually show different types of populations, without any assumptions of the statistical distribution. The spacing’s between the different parts of the box help indicate variance, skew and identify outliers.

Box plots can be drawn either horizontally or vertically.

5.3. Data Analyses

5.3.1. Factor Analyses

The factor analysis is used as a statistical data reduction technique. It explains variability among observed random variables in terms inn the connection of fewer unobserved random variables. This random Variables are called factors. The observed variables are modelled as linear combinations of the factors, plus "error" terms. (Wikipedia, http://en.wikipedia.org/wiki/Factor_analysis, accessed in 07/2007, Darlington, et al.1973).

Statistical methods are normally used to study the relationship between independent and dependent variables. The factor analysis has a different approach. It is used to study the patterns of relationships among many dependent variables The goal of the analyses is to discover something about the nature of the independent variables that affect them. This is also done when those independent variables were not measured directly. The inferred independent variables are called factors. A typical factor analysis suggests answers following four major questions:

1. How many different factors are needed to explain the pattern of relationships among these variables?

2. What is the nature of those factors?

3. How well do the hypothesized factors explain the observed data?

4. How much purely random or unique variance does each observed variable include?

5.3.2. Principal Component Analysis (PCA)

The PCA is a technique which is used to reduce multidimensional data sets to lower dimensions for the analysis. Depending on the field of application, it can also be named the discrete Karhunen-Loève transform, the Hotelling transform or proper orthogonal decomposition (POD). (Wikipedia, http://en.wikipedia.org/wiki/Factor_analysis, accessed in 07/2007, Darlington, et al.1973).

The goal of the categorical principal components analysis (CATPCA) is to reduce an original set of variables into a smaller set of uncorrelated components that represent most of the information found in the original variables. The technique is most useful when a large number of variables prohibit effective interpretation of the relationships between objects (subjects and units). Degerman et al. 2007 used the PCA analyses to find correlations between impact Variables.

5.3.3. Discriminant Analysis

The discriminant analysis is a technique to classify a set of observed variables into predefined classes. The main goal is to determine the class of an observation based on a set of variables known as predictors or input variables.

The methods used are: Multiple Discriminant Analysis, Fisher's Linear Discriminant Analysis, and K-Nearest Neighbours Discriminant Analysis. (ESO 1999)

5.3.4. Cluster Analyses

Cluster analyses areused to classify objects into different groups. Described more precisely its used for the partitioning of a data set into subsets (clusters) The data in each subset should share ideally some common characteristics. Data clustering is a common technique for statistical data analysis, which is used in many fields, including machine learning, data mining, pattern recognition, image analysis and bioinformatics. (Wikipedia, http://en.wikipedia.org/wiki/Cluster_analysis accessed in 07/2007). Melcher et al. (2007) used the cluster analyses for a spatially based method to assess the ecological status of European fish assemblage types.

5.3.5. Spearman's Rank Correlation

The Spearman's Rank Correlation is a technique used to test the direction and strength of the relationship between variables. It is a method to show if a variable has an effect on another one, (RevisionNotes.Co.Uk, http://www.revision-notes.co.uk/revision/181.html, accessed in 01/2008, accessed in 07/2007).

Correlation between Variables:

-1 → perfect negative correlation

In between -1 and -0.5 → strong negative correlation

In between -0.5 and 0, → weak negative correlation

0 → there is no correlation

In between 0 and 0.5, → weak positive correlation

In between 0.5 and 1, → strong positive correlation

1, there is a perfect positive correlation

5.4. SPSS

5.4.1. Description of SPSS:

The Statistical Package for the Social Sciences, (SPSS) for Windows is a powerful statistical analysis program for the MS Windows computer system.

SPSS is one of the widely used programs for statistical analyses in social and business science. The first version was released in 1968 and the up do date version is version 15. It is intensively used by health researchers, market researchers, survey companies, government, education researchers at university and schools and others. In addition to the statistical analysis, data management (case selection, file reshaping, creating derived data) and data documentation (a metadata dictionary is stored with the data) are features of the base software.

The features of SPSS are easily accessible with pull-down menus and/or can be programmed with a proprietary 4GL command syntax language. This command syntax Language has the benefits of the easy reproducibility for the handling of complex data analyses and manipulations, The pull-down menu interface can also generate a command syntax, but therefore the default settings have to be changed to make them visible for user.

To write command language subroutines a "macro" language can be used. Also a Python programmability extension can access the information in the data dictionary and dynamically build command syntax programs. The previous mentioned programmability extension, implemented in SPSS 14, replaced the less functional SAX Basic "scripts" for most purposes. From this second to the last version the programme can be driven externally by a Python or a VB.NET program using supplied "plug-ins".

SPSS places the importance on data processing, data types, the internal structure of files and matching files. Together it consiterably simplifies programming. Datasets used in SPSS have a 2-dimensional Table structure where the x-axes rows can typically represent cases (e.g. individuals, households) and the y-axes columns represent measurements (e.g. age, sex or household income, smokers non smokers).

Three data types can be defined, metric, ordinal and categorical. All data processing occurs sequentially case-by-case through the file. Files can be matched one-to-one and one-to-many, but not many-to-many.

Different versions of SPSS are available for this four operating systems: Windows, Mac OS X and Unix. Only the Windows version is updated more often, and has more features, than the versions.

SPSS can also read and write data from other statistics packages, spreadsheets, databases, external relational database Tables via ODBC and SQL and ASCII text files.

The proprietary SPSS file formats are: .spo and .sav.

Also a stand alone reader (draft viewer) is provided, for which, in addition to the in-package viewer the output can be produced in text form only, captured as data text, tabdelimited text (e.g. HTML, XML) dataset and as a variety of graphic image formats (e.g. JPEG, PNG, BMP and EMF).

5.4.2. Features of SPSS:

● Descriptive statistics:

Cross tabulation, Frequencies, Descriptives, Explore, Descriptive Ratio Statistics

● Bivariate statistics:

Means, t-test, ANOVA, Correlation (bivariate, partial, distances), nonparametric tests

● Prediction for numerical outcomes:

Linear regression

● Prediction for identifying groups:

Factor analysis, cluster analysis (two-step, K-means, hierarchical)

5.4.3. Add - On Modules Available

● SPSS Programmability Extension (added in version 14):

Allows Python programming control of SPSS.

● SPSS Data Validation (added in version 14):

Allows programming of logical checks and reporting of suspicious values.

● SPSS Regression Models:

Logistic regression, ordinal regression, multinomial logistic regression, and mixed models (multilevel models).

● SPSS Advanced Models:

Multivariate GLM and repeated measures ANOVA (removed from base system in version 14).

● SPSS Classification Trees: Creates classification and decision trees for identifying groups and predicting behaviour.

● SPSS Tables:

Allows user-defined control of output for reports.

● SPSS Exact Tests: Allows statistical testing on small samples.

● SPSS Categories

● SPSS Trends™

SPSS Conjoint, SPSS Missing Value Analysis: Simple regression-based imputation.

● SPSS Map

● SPSS Complex Samples (added in Version 12):Adjusts for stratification and clustering and other sample selection biases.

● SPSS Server is a version of SPSS with a client/server architecture. It has some features not available in the desktop version, one example is scoring functions.

5.5. Environmental Variables and Sampling Methods

Fish sampling

To obtain the necessary fish data for the index electric fishing can be used. These standardised electric fishing procedures are described in the CEN directive, "Water Analysis – Fishing with Electricity (EN 14011; CEN, 2003) for wad able and non-wad able rivers.

The electric fishing methods differ depending upon the water depth and wetted width of the sampling site. The selection of waveform (DC (Direct Current) or PDC (Pulsed Direct Current)), depends on: 1. the conductivity of the water, 2. the dimensions of the water body and 3. the fish species to be expected. Alternating Current (AC), which is harmful for fish should not be used.

The fishing procedure is summarised below, separately for wad able and non-wad able rivers. In both cases, fishing equipment must be suitable to sample small / younger individuals.

In the following Table 4, the river width corresponds with wetted width according to the CEN-standard, the main purpose of the standardised sampling procedure is to record information concerning fish composition and abundance; therefore, no sampling period is defined (according to CEN). However, the FAME project agreed on a sampling period of late summer/early autumn.

The only exception was made for non-permanent Mediterranean rivers where samples in the spring may be more appropriate.

Table 4. CEN Standard (CEN directive 2003)

Waveform selection:	DC or PDC
Number of anodes:	One anode per 5 m width
Number of hand-netters:	Each anode followed by 1 or 2 hand-netters (mesh size of 6 mm maximum) and 1 suitable vessel for holding fish.
Number of runs:	One run
Time of the day:	Daylight hours
Fishing length:	10 - 20 times the wetted width, with a minimum length of 100 m
Fished area:	river width <15 m: The whole site surface river width >15 m: Several separated sampling areas are selected and prospected within a sampling site, with a minimum of 1000 m² (partial sampling method)
Fishing direction:	Upstream
Movement:	Slowly, covering the habitat with a sweeping movement of the anodes and attempt to draw fish out of hiding.
Stop nets:	Used if necessary and feasible

Waveform selection:	DC or PDC
Number of anodes:	Depending on boat configuration
Number of runs:	One run
Time of the day:	Daylight hours
Fishing length:	10 -20 times the wetted width, with a minimum length of 100 m
Fished area:	Both banks of the river or a number of sub-samples proportional to the diversity of the habitats present with a minimum of 1000 m² (partial sampling method)
Fishing direction:	Normal flow: downstream in such a manner as to facilitate good coverage of the habitat, especially where weed beds are present or hiding places of any kind are likely to conceal fish High flow: upstream Low flow: not necessary to match boat movement to water flow, and the boat can be controlled by ropes from the bank side if required
Movement:	Slowly, covering the habitat with a sweeping movement of the anodes or drifting with the boom along selected habitats and attempting to draw fish out of hiding.
Stop net	Used if necessary and feasible

The minimum river length to be sampled, because of the:

● variability of habitats and fish communities within rivers sections and

● in order to ensure accurate characterisation of a fish community,

electric fishing at a given site must be conducted over a river length of:

10 to 20 times the river width, with a minimum length of 100 m.

Large and Shallow Rivers (width >15 m and water depth <70 cm):

The length of the sampling site (station) is also calculated as 10 to 20 times the river width. Fishing of longer river sections should be avoided as some metrics referring to the number of species caught (e.g. number of rheophilic species) might be biased due to over sampling.

In wadable rivers as a general guide one anode per 5 m width should be appropriate. The operators should generally fish upstream so that the from the wading shoes dispersed soil of the river base does not affect the fishing efficiency.

They operators should move slowly, covering the hole habitat with a sweeping movement of the anodes and attempt to draw fish out of hiding. To aid effective fish capture in fast flowing water the catching nets should be held in the wake of the anode. Each anode is generally followed by one or two hand-netters (hand net: mesh size of 6 mm maximum) and one suitable vessel for transporting fish.

Large rivers (depth > 0.7 m)

The large variety of habitats in large rivers (depth > 0.7 m) makes the analyses of the entire area almost impossible. Therefore, a partial sampling procedure should be applied to cover all types of habitats to obtain a representative sample of the site. Qualitative and semi-quantitative information can be gained by using conventional electric fishing methods with hand held electrodes in the river margins and delimited areas of habitat. The capture efficiency can be improved by using electric fishing boats with booms which increase the size of the effective electric field relative to the area being fished by increasing the number of catching electrodes. Depending on the water conductivity, the current demands of multiple electrodes can be high and large generators may be needed.

5.6. Approaches of the Pressure Evaluation

The following three approaches of a index development were discussed by experts at the second EFI+ Workshop in Lisbon 16-19, October 2007 and will be also discussed in this thesis.

FAME Approach

Principle: Arithmetic mean

Waterquality	Morphology	Continuum	Hydrology
1	2	1	5

Calculation: (1+2+1+5)/4=2.25

Index Example: 2.25

The limitation of this approach is that only an average is consitered and that single pressures are underrepresented.

Degraded Approach

Principle: Avarage of all values worse than 2

Waterquality	Morphology	Continuum	Hydrology
1	2	1	5

Calculation: (2+5)/2=3.5

Index Example: 3.5

The limitation of this approach is that unimpacted pressure variables are not considered.

Worst Case Approach

Principle One out –all out

Waterquality	Morphology	Continuum	Hydrology
1	2	1	5

Calculation: “worst value” is always taken: e.g.: If 3 is the highest count,3 is taken

e.g. if 5 is the highest count, 5 is taken Therefore:

Index Example: 5

The limitation of this approach is that the worst variable always will influence the whole classification.

6. Results

6.1. Overall Completeness of Data

The following chapter gives a short overview of the overall completeness of all sampled and calculated variables. In total 29509 sites of 15 countries were measured. Counts marked in bolded counts represent pressures with critical number of sites without valid information on specific pressures. For the analyses of completeness of data the dataset-version from December 2007 was used because at this time the final version was not available.

A detailed description of the overall completeness of data per country will is given in the annex.

6.1.1. Completeness of Environmental Variables

The overall completeness for the dataset with 29509 sites of all environmental variables is 85.48%. The completeness of the variables Water_source_type, Valley_slope and Floodplain lies under 90 % which explains the relative low arithmetic mean of the environmental variables data set.

Table 5: Completeness of environmental variables

Variable	Completeness in %
Altitude	100
Lakes_upstream	92.5
Distance_from_source	99.1
Flow_regime	96.2
Geological_typology	97.6
Water_source_type	**76.8**
Actual_river_slope	99,0
Valley_slope	**65.3**
Size_of_catchment	99.6
Catchment_name	100.0
Floodplain	91.3
Valley_form	94.3
Geomorph_river_type	91.8
Natural_sediment	92.3

6.1.2. Completeness of Variables Describing the Sampling Strategy

The overall completeness for all variables describing the sampling site is 95.9%

Table 6 Completeness of variables describing the sampling strategy

Variable	Completeness in %
Sampling_strategy	93.9
Method	97.9
Fished_area	95.9

6.1.3. Completeness of Variables Describing the Location, Name of Site and Date of Fishing

The overall completeness for the variables describing the location, Name of the site and date of fishing is 100 %. Especially for the Site code it is important to have complete data from 100% because this is important for further calculations.

Table 7: Completeness of variables describing the location, name of site and date of fishing

Variable	Completeness in %
Site_code	100.0
Date	100.0
Latitude	100.0
Longitude	100.0
River_name	100.0
Site_name	100.0
Fides_site	100.0

6.1.4. Completeness of Variables Describing Pressures

The overall completeness for the hydrological, morphological, water quality and connectivity pressures types are displayed in the column “Arithmetic Mean Pressure Type”. The arithmetic mean for all pressure types is very low because the low completeness of the pressure type connectivity influences completeness of all other pressure types.

Table 8: Completeness of variables describing pressures

Pressurestype:	**Variable:**	**Completeness in %.**	**Arithmetic Mean Pressure Type:**
Hydrological	Impoundment	95.1	92.7
	Hydropeaking	98.8	
	Colinear_connected_reservoir	91.2	
	Hydro_mod	97.8	
	Temperature_impact	**86.8**	
	Velocity_increase	**84.2**	
	Reservoir_flushing	98.4	
	Sedimentation	**89.6**	
Morphological	Channelisation	93.8	88.9
	Cross_sec	92.9	
	Instream_habitat	92.6	
	Riparian_vegetation	**84.4**	
	Embankment	91.6	
	Floodprotection	92.1	
	Floodplain_site	**75.5**	
Water Quality	Toxic_substances	**88.8**	88.4
	Acidification	97.4	
	Water_quality_index	**83.0**	
	Water_quality_name	83.1	
	Eutrophication	93.7	
	Organic_pollution	**88.8**	
	Organic_siltation	**84.3**	
	Navigation	**88.2**	
Connectivity	Barriers_catchment_down	98.5	74.8
	Barriers_river_segment_up	98.3	
	Barriers_river_segment_down	98.3	
	Barriers_number_river_segment_up	**82.7**	
	Barriers_number_river_segment_down	**81.3**	
	Barriers_distance_river_segment_up	**34.0**	
	Barriers_distance_river_segment_down	**30.4**	
	Arithmetic Mean All PressureTypes:		86.5

6.1.5. Basic Completeness of Pressures Variables for Each Country

The completeness of the anthropogenic pressures was evaluated for the dataset of 13085 sites. The detailed tables of all pressures with the missing and and other existing counts, are presented in the annexes.

The following table gives an short overview about the missing data of each country for each pressure:

"Sum" describes the number of sites with valid values for each variable. "Miss" gives information on missing data for each variable. Counts marked in bold black represent pressures with critical number of sites without valid information on specific pressures. All variables used in the following table are the same as used in the previous chapter. They are only recoded to numerical values, which was necessary because, some statistical calculations are not possible for categorical variables.

Remarkable is the Switzerland has an high amount of missing variables for morphological and water quality pressures. Finland is the country with the most missing variables for all pressure types. Italy's missing values are concentrated on morphological pressures, which is the same fact for United Kingdom.

Table 9. Overview on completeness of datasets for each pressure variable and country.

Country	AT		CH		DE		ES		FI		FR		HU		IT		LT		NL		PL		PT		RO		SE		UK	
	Sum	Miss	Sum	Miss	Sum	Miss	Sum	Miss	Sum	Miss	Sum	Miss	Sum	Miss	Sum	Miss	Sum	Miss	Sum	Miss	Sum	Miss	Sum	Miss	Sum	Miss	Sum	Miss	Sum	Miss
H_hydrop	918	4	717	0	785	0	3147	0	**304**	**226**	1115	30	193	0	652	0	114	1	182	0	916	3	923	0	263	0	605	0	1987	0
H_waterabstr	915	7	717	0	785	0	**2134**	**1013**	**304**	**226**	1082	63	193	0	564	88	115	0	182	0	919	0	923	0	263	0	605	0	1987	0
H_hydromod	922	0	717	0	785	0	3147	0	**304**	**226**	1093	52	193	0	563	89	95	20	182	0	915	4	923	0	263	0	605	0	1980	7
H_tempimp	898	24	717	0	785	0	3147	0	**304**	**226**	1079	66	193	0	583	69	97	18	182	0	913	6	923	0	263	0	605	0	**0**	**1987**
H_resflush	889	33	717	0	785	0	3147	0	**304**	**226**	1118	27	193	0	652	0	115	0	182	0	918	1	923	0	263	0	605	0	1987	0
H_imp	921	1	717	0	785	0	3147	0	**304**	**226**	1107	38	193	0	652	0	115	0	182	0	919	0	923	0	263	0	605	0	**1264**	**723**
H_veloincr	920	2	717	0	785	0	2884	263	**304**	**226**	1102	43	193	0	569	83	91	24	182	0	916	3	923	0	263	0	605	0	**0**	**1987**
M_sed	922	0	**82**	**635**	777	8	**2358**	**789**	**304**	**226**	1101	44	193	0	652	0	106	9	182	0	915	4	923	0	263	0	605	0	1980	7
M_channel	922	0	717	0	785	0	2884	263	**281**	**249**	1116	29	193	0	**546**	**106**	115	0	182	0	919	0	923	0	263	0	605	0	**1645**	**342**
M_crossec	922	0	**557**	**160**	785	0	2883	264	**281**	**249**	1111	34	193	0	**551**	**101**	115	0	182	0	919	0	923	0	263	0	605	0	**1683**	**304**
M_instrhab	922	0	**534**	**183**	785	0	2881	266	**281**	**249**	1110	35	193	0	**547**	**105**	115	0	182	0	919	0	923	0	263	0	605	0	**1683**	**304**
M_ripveg	922	0	717	0	785	0	2880	267	**310**	**220**	1113	32	193	0	**547**	**105**	115	0	182	0	918	1	923	0	263	0	605	0	**0**	**1987**
M_embank	922	0	**552**	**165**	785	0	2884	263	**310**	**220**	1115	30	193	0	**548**	**104**	115	0	182	0	918	1	923	0	263	0	605	0	**1394**	**593**
M_floodpro	922	0	717	0	785	0	2880	267	**310**	**220**	1117	28	193	0	**544**	**108**	115	0	182	0	916	3	923	0	263	0	605	0	**1383**	**604**
M_remfloodpl	922	0	**316**	**401**	733	52	**2312**	**835**	**310**	**220**	1080	65	193	0	652	0	6	109	182	0	**284**	**635**	923	0	254	9	605	0	**0**	**1987**
C_B_c_do	922	0	717	0	781	4	3147	0	**310**	**220**	1105	40	193	0	652	0	115	0	182	0	919	0	923	0	263	0	605	0	1987	0
C_B_s_up	922	0	694	23	781	4	3143	4	**310**	**220**	1075	70	193	0	652	0	115	0	182	0	919	0	923	0	263	0	605	0	1987	0
C_B_s_do	922	0	694	23	781	4	3132	15	**310**	**220**	1077	68	193	0	652	0	115	0	182	0	919	0	923	0	263	0	605	0	1987	0
C_Bn_sup	922	0	**378**	**339**	784	1	3143	4	**397**	**133**	**446**	**699**	193	0	651	1	51	64	182	0	919	0	923	0	263	0	605	0	1987	0
C_Bn_sdo	922	0	**344**	**373**	784	1	3123	24	**397**	**133**	**440**	**705**	193	0	652	0	43	72	182	0	919	0	923	0	263	0	605	0	1987	0
C_Bd_sup	922	0	**375**	**342**	**632**	**153**	**938**	**2209**	**397**	**133**	**473**	**672**	**61**	**132**	651	1	51	64	182	0	**288**	**631**	**214**	**709**	263	0	**180**	**425**	**65**	**1922**
C_Bd_sdo	922	0	**344**	**373**	**631**	**154**	**868**	**2279**	**397**	**133**	**464**	**681**	**61**	**132**	646	6	43	72	182	0	**306**	**613**	**225**	**698**	263	0	**151**	**454**	**65**	**1922**
W_toxic	922	0	**79**	**638**	762	23	3129	18	**310**	**220**	892	253	193	0	**384**	**268**	115	0	182	0	914	5	923	0	263	0	605	0	1980	7
W_acid	922	0	717	0	785	0	3147	0	**310**	**220**	1059	86	193	0	652	0	115	0	182	0	915	4	923	0	263	0	605	0	1980	7
W_index	922	0	713	4	785	0	2751	396	**310**	**220**	684	461	193	0	**492**	**160**	115	0	182	0	912	7	923	0	263	0	566	39	1972	15
W_eutroph	922	0	**79**	**638**	785	0	3099	48	**310**	**220**	1093	52	193	0	628	24	115	0	182	0	917	2	923	0	263	0	605	0	1980	7
W_opoll	916	6	**241**	**476**	785	0	3054	93	**310**	**220**	**900**	**245**	193	0	578	74	115	0	182	0	917	2	923	0	263	0	605	0	1980	7
W_osilt	922	0	**241**	**476**	785	0	3112	35	**310**	**220**	1086	59	193	0	522	130	114	1	182	0	917	2	923	0	263	0	605	0	**0**	**1987**
O_nav	922	0	717	0	785	0	3147	0	**310**	**220**	1127	18	193	0	652	0	115	0	182	0	919	0	923	0	263	0	605	0	**0**	**1987**
O_collconn	922	0	**0**	**717**	785	0	2786	361	**310**	**220**	1109	36	193	0	652	0	115	0	182	0	918	1	923	0	263	0	605	0	1987	0

6.2. Selection of Pressures for Pressure Analyses

To take into account standards as the EN 14962 („Water quality – guidance on the scope and selection of fish sampling methods“, CEN 2004), the total dataset (N=29509, has been filtered before analysing pressures as follows:

Only sites were taken with complete information on Run1 in table catch and environmental variables like altitude, actual slope, wetted width and fished area. If the information on this variable were not complete, the sites were drop out.

With this step the dataset was reduced from 29509 sites to 27338 sites. That is equal to a loss of 2171 sites or 7.4 % of the data available.

6.2.1. Amount of Sampling per Site, Sampling Strategy of different Countries

As the participating countries of this project have different sampling strategies and also the amount of money reserved for scientific research programmes is restricted, the amount of samples on each site is different. The table 10 gives a short overview for each country how often a sample was taken on each site and the table 11 explains the country abbreviations used for partners of the project:

Table 10: Abbreviations of countries

Abbreviation	**SE**	**FR**	**NL**	**DE**	**IT**	**UK**	**CH**	**AT**
Country	Sweden	France	Netherlands	Germany	Italy	United Kingdom	Switzerland	Austria
Abbreviation	**RO**	**ES**	**LT**	**PL**	**FI**	**HU**	**PT**	
Country	Romania	Spain	Lithuania	Poland	Finland	Hungary	Portugal	

Table 11: Amount of sampling per site

Country	SE	FR	NL	DE	IT	UK	CH	AT	RO	ES	LT	PL	FI	HU	PT
Amount of samples on each site	9.2	5.7	4.3	2.3	1.8	1.6	1.4	1.2	1.2	1.2	1.0	1.0	1.0	1.0	1.0

Due to the fact that table fishing occasion contains a lot of replicates for each fishing site, a solution which fishing occasions is taken for the analyses had to be found.

Therefore, from all sites only the youngest site (youngest date) was filtered out. Finally, for each site only one sampling date was given so that the weighting of the data is equal. After that procedure from the in the previous mentioned 27338 sites only 13085 were left over. That is equal to an loss of 14525 sites or 52,1% of the data.

6.3. Distribution of Sampling Sites

The following table 12 and figure 10 gives an over few of the amount of sampling sites. Spain, France and United Kingdom have the highest amount of sampled sites.

Table 12: Distribution of sampling sites

Distribution of Sampling Sites																
	AT	**CH**	**DE**	**ES**	**FI**	**FR**	**HU**	**IT**	**LT**	**NL**	**PL**	**PT**	**RO**	**SE**	**UK**	**Total**
Nr.	922	717	785	3147	530	1145	193	652	115	182	919	923	263	605	1987	13085
%	7.0	5.5	6.0	24.1	4.1	8.8	1.5	5.0	0.9	1.4	7.0	7.1	2.0	4.6	15.2	100

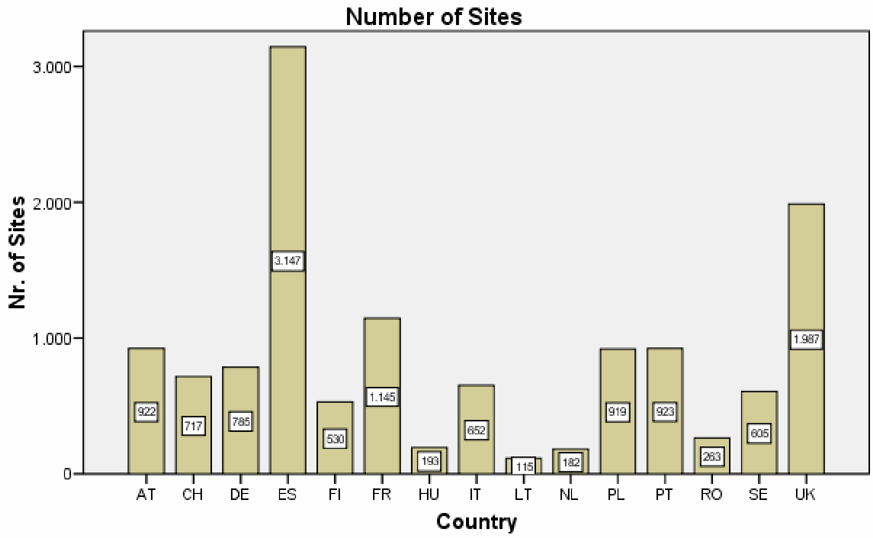

Figure 10: Distribution of sampling sites

6.4. Analyses of Pressures

The basis for the following analyses is the previous filtered dataset with the 13085 sites.

6.4.1. Number of EFI+ Sites per Country

The following table 13 is a short summary of number of sites per country. The number of total sites in table site of EFI+ and number of total fishing occasions in Table fishing occasion of EFI+ are compared.

The variants of sites, Max Data loss Consensus1, Consensus 2 and Min Data Loss are only given in this table for completation and will be explained later in the chapter: Pressure Index Development.

Table 13: Comparison between the number of sites and fishing occasions in EFI+ and finally selected sites (after filtering) for pressure analyses.

Country	Nr. of sites selected for pressure analysis	Nr of sites in Table site/ EFI+	Nr. of fishing occasion in Table fishing occasion/ EFI+	Max Data Loss	Consensus 1	Consensus 2	Min Data Loss
AT	922	938	1172	847	873	879	879
CH	717	717	969	0	0	70	428
DE	785	803	1817	741	741	764	772
ES	3147	4239	5189	1650	1596	1596	1613
FI	530	530	530	257	257	257	257
FR	1145	1145	6570	0	594	874	808
HU	193	193	193	166	166	166	166
IT	652	652	1152	303	315	468	469
LT	115	115	130	54	83	83	90
NL	182	182	790	182	182	182	182
PL	919	919	978	884	888	893	893
PT	923	923	923	923	923	923	923
RO	263	263	323	262	262	262	262
SE	605	615	5652	602	602	602	602
UK	1987	1987	3162	0	1154	1154	1154
Total	13085	14221	29550	6871	8636	9173	9498

6.4.2. Index Scenarios

For the comparison of the Index, different variants with different variables and completeness of data where done. The variables were selected by expert judgment. From the worst variant to the best variant, the amount of variables is declining but therefore also the amount of the numbers of retained sites is rising. The exceptions are the Variants Consensus 2 and Min Data Loss where the amount of variables is the same.

The different scenarios are as follows:

1. Variant: 13085 Sites

This dataset is explained in the chapter: 6.2 Selection of Pressures for Pressure Analyses

2. Variant, "Maximum Data Loss":

After elimination of the missing data for this variant only 6871 sites were left

For the countries Switzerland, France and United Kingdom no data were left because of their high amount of missing data. In total 34 variables were used.

3. Variant, "Consensus 1":

This variant is the 2^{nd} worst variant. In total 8636 are left and 30 variables were used.

Switzerland is still an exception because of there high amount of no data.

1. Variant, "Consensus 2":

This variant is the 2^{nd} best variant. 9173 sites are left and 26 variables were used.

For all 15 countries an Index can be calculated.

1. Variant, "Minimum Data Loss":

This variant is the best variant. In total 9173 sites are left and again 26 variables but different to the variables of the variant, "Consensus 2" were used.

6.4.3. Pressure Groups

As a basis for pressure analyses, the following five pressure groups have been defined:

1. hydrological pressures, (7 variables)
2. morphological pressures (9 variables)
3. connectivity pressures, (7 variables)
4. water quality pressures, (6 variables)
5. and other pressures which do not fit into previous mentioned groups. (2 variables)

The variables give information on human pressures through direct effects or by describing them indirectly (descriptors).

The definition of the variables was done by expert judgement during preliminary consortium meetings and finally fixed at the EFI+ Kick-off meeting in January 2007 in Bratislava. This was important to guarantee that each country with its typical geographic region can be considered with its specific pressures in the EFI+ database.

For further analyses, each variable has been transform into “numeric values” according to a scheme from 1-5, where 1 is the best and 5 is the worst.

The coding scheme has been well consitered and was applied depending on the variable in the same way:

“No” = 1,

“Slight” = 2,

“Weak” or “Intermediate” = 3

“High or Intermediate” =4

“High” or “Strong” = 5

Variables which describe diametric effects (e.g. variable remaining floodplain in the group of morphological pressures) or consist of more than 4 or 5 categories have been coded in a special manner.

6.4.4. Environmental Variability

The following analyses are important to get an overview about the spatial distribution of sites and to also find sites with similar environmental conditions.

Distribution of Altitude

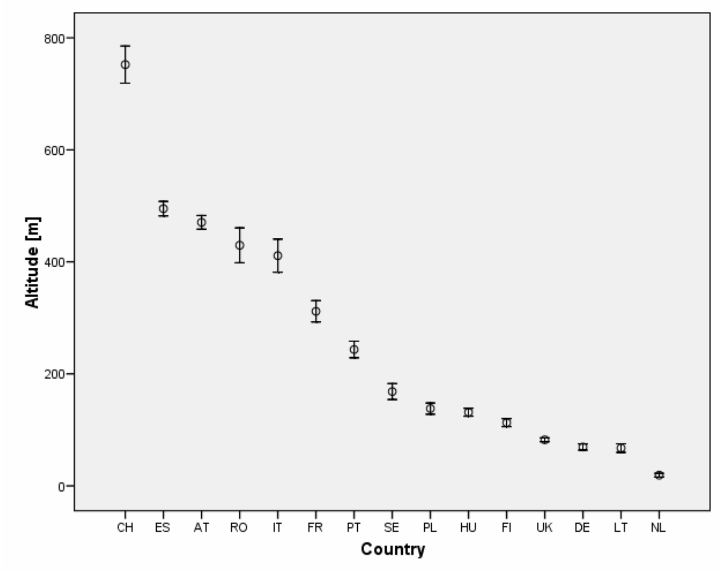

Figure 11: Distribution altitudes per country rank (error bar 95% confidential interval of altitude)

In the figure 11 the distribution of the altitude per site is shown. The highest sites are located in Switzerland (the mean above see level is around 700 m) and the lowest site are in the Netherlands.

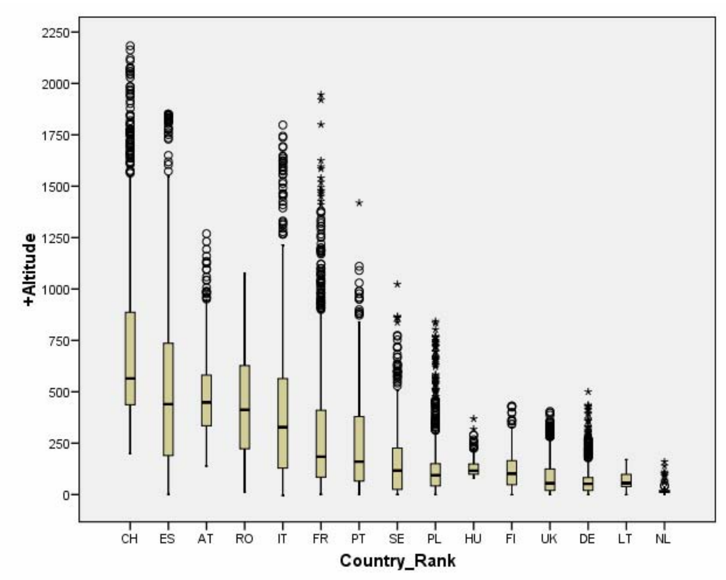

Figure 12: Distribution altitudes per country rank

In the figure 12 it can be seen that the median above see level from Switzerland went down to 520 m. The reason for that can be found in the high amount of outliers and extreme values.

Distribution of Actual River Slope:

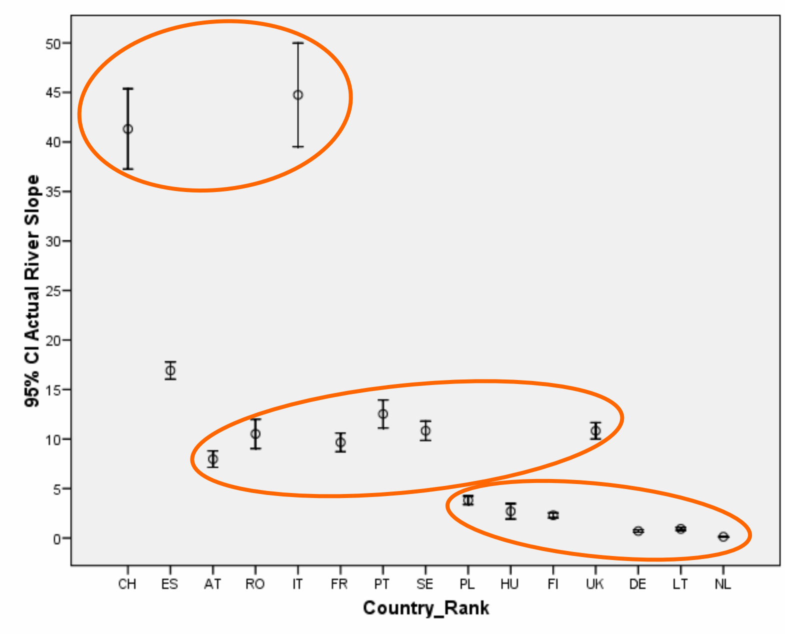

Figure 13: Distribution actual river slope per country rank (error bar 95% confidential interval of actual river slope)

The highest actual river slopes from all project partners are located in Switzerland and Italy. Two other groups with the same slope can be also observed. Only Spain doesn't fit into orange ellipse market groups.

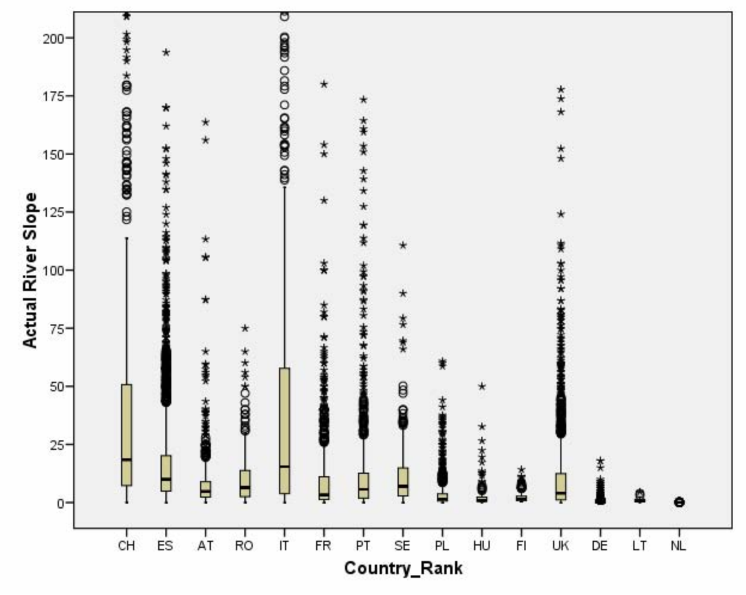

Figure 14: Distribution actual river slope per country rank

The median of the box plots show a lower value for Switzerland and Italy which can be again explained by high amount of outliers and extreme values.

Distribution of Wetted Width:

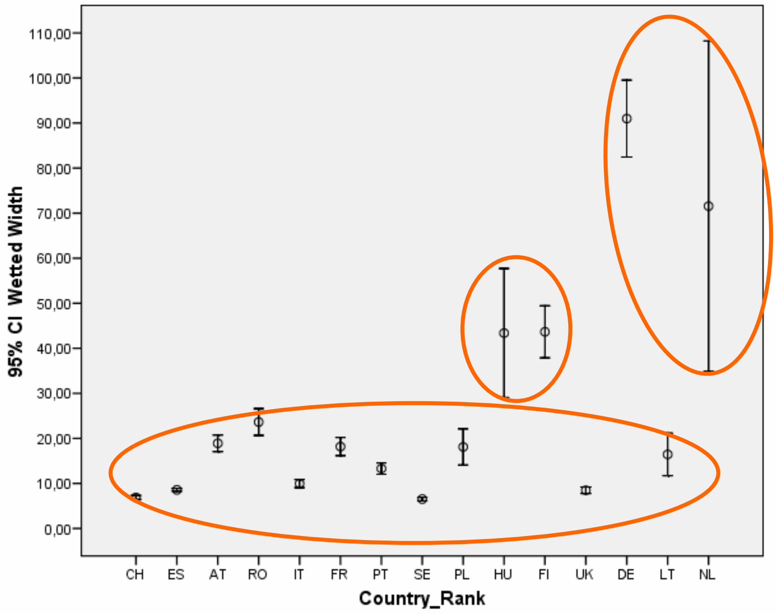

Figure 15: Distribution wetted width per country rank (error bar 95% confidential interval of wetted width)

In the figure 15, the error bar can be again differentiated in three different groups. Interesting is the highest group with the Netherlands. Following from the error bar it seems that in the low river systems near the ocean have a very broad wetted width.

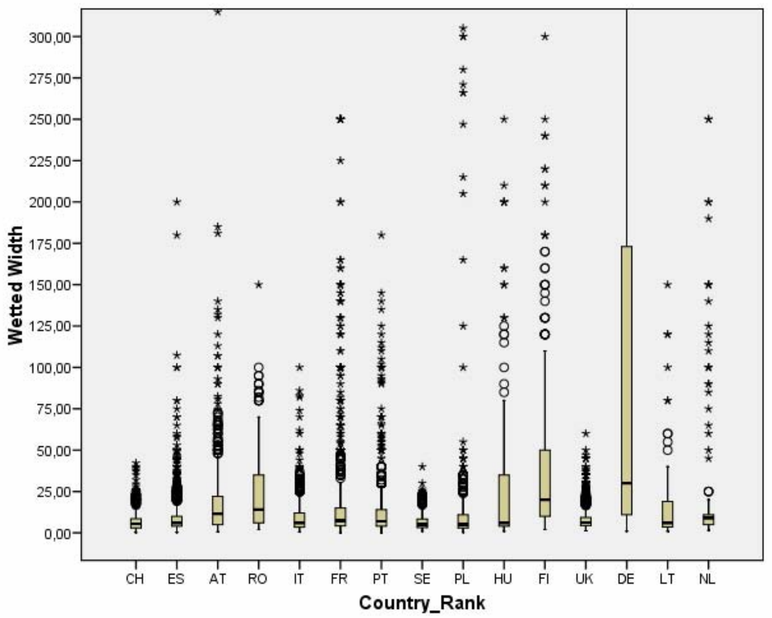

Figure 16: Distribution wetted width per country rank

Distribution of Size of Catchment:

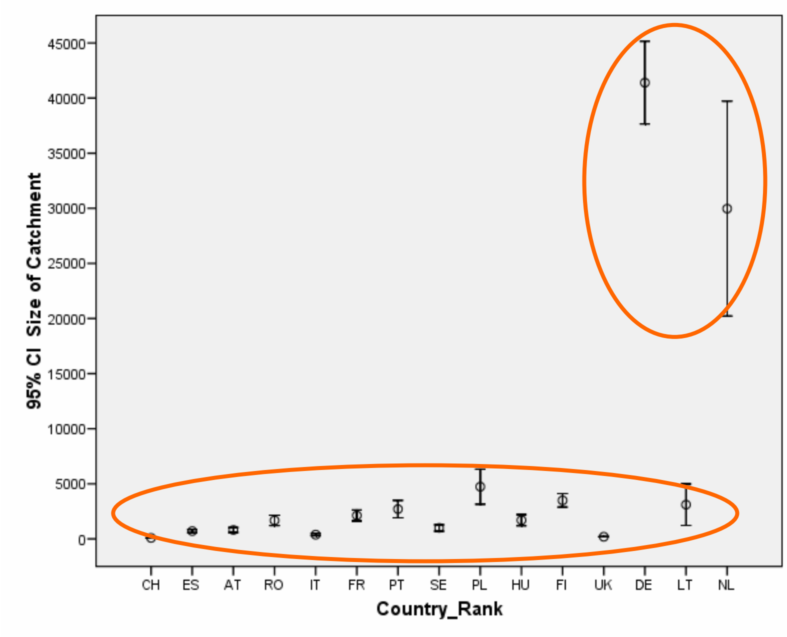

Figure 17: Distribution size of catchment per country rank (error bar 95% confidential interval of size of catchment)

For the size of catchment two groups can be differentiated. The first group contains catchments from 0 to 500 km^2. The second group contains Lithuania with 4500 km^2 and Netherlands 3000 km^2.

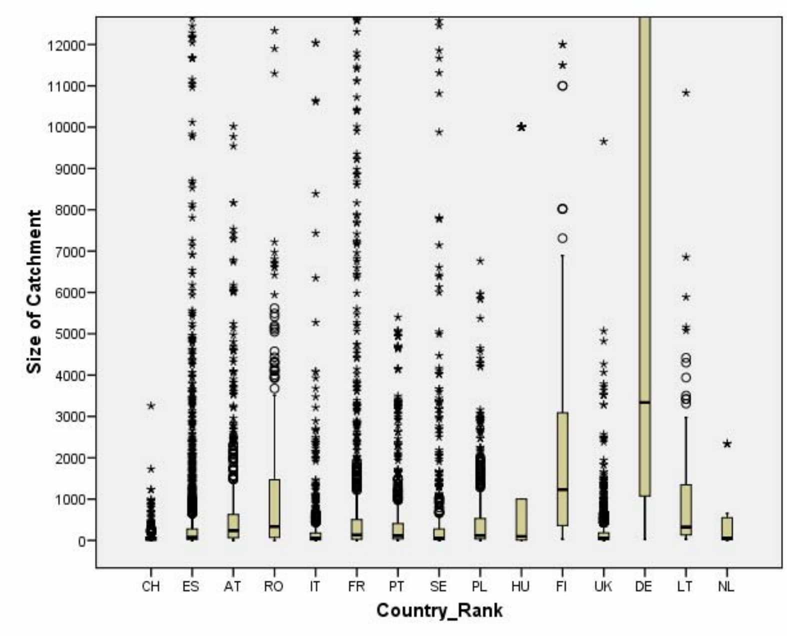

Figure 18: Distribution size of catchment per country rank

Factor analysis

The following variables which are consitered in the “Table Site” to be potential descriptors of environmental variability have been selected:

1. altitude
2. actual river slope (taken after correlation was checked with variable valley slope)
3. wetted width
4. size of catchment
5. latitude and longitude.

Later on a factor analysis with these input variables was done.

The output of this analyses are: 3 different factors and around 79,5% explained variance.

The factors were summed as:

1. FAC1_catch (contains variables: size_of_catchment and wetted_width),
2. FAC2_geo (contains variables latitude and longitude) and finally
3. FAC3_alt (contains variables altitude and actual river slope).

Table 14: Total variance for 6 environmental input variables

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	2.139	35.651	35.651	2.139	35.651	35.651	1.704	28.404	28.404
2	1.444	24.064	59.715	1.444	24.064	59.715	1.586	26.439	54.843
3	1.185	19.748	**79.464**	1.185	19.748	79.464	1.477	24.620	79.464
4	.620	10.334	89.798						
5	.317	5.291	95.089						
6	.295	4.911	100.000						

Table 15: Rotated component matrix, describing he composition of factors FAC1_catch, Fac2_geo and FAC3_alt.

Component	1 FAC1_catch	2 Fac2_geo	3 Fac3_alt
Size_of_catchment	**.919**		
Wetted_width	**.917**		
Longitude		**.898**	
Latitude		**.858**	
Actual_river_slope			**.837**
Altitude			**.829**

Extraction Method: Principal Component Analysis. Rotation Method:

Cluster analysis

For the cluster analyses 3 factors were used as input data for the K-means method. The output were 6 clusters representing different types of environmental variability.

Table 16: 6 final cluster centers, describing the composition out of the 3 different factors, representing types of environmental variability.

	Cluster					
	1	2	3	4	5	6
FAC1_catch	-,10910	3.33127	22.48782	-.01040	6.82281	-.21245
FAC2_geo	-.99866	.43188	-.88885	.30408	.08138	.62153
FAC3_alt	.10376	-.40471	.79682	3.34955	-.26195	-.31118

Due to the fact that cluster number 3 contained only 3 sites (located in the Netherlands at river Meuse with a very large catchment). this cluster has been combined with cluster 6. The cluster-order has also been re-structured so that, it represents 5 different river types of environmental variability.

Table 17: Types of environmental variability after reduction from 6 to 5 clusters.

		Frequency	Percent	Valid %	Cumulative %
Valid	**1**	**557**	4.3	4.4	4.4
	2	**4703**	35.9	36.8	41.2
	3	**7056**	53.9	55.3	96.5
	4	**309**	2.4	2.4	98.9
	5	**138**	1.1	1.1	100.0
	Total	**12763**	97.5	100.0	
Missing	System	322	2.5		
Total		13085	100.0		

Discriminant analysis

To test the output of cluster analysis a discriminant analysis has been done. The variance of discriminant analysis was around 83,3%. Due to a better variability within the 5 river types it was decided to take the 5 groups classified by discriminant analyses for all further analyses.

Table 18: Comparison of cluster 5 (counts after cluster analyses and optimizing) and discrim 5 (counts after discriminant analysis).

Count	cluster5					Total
discrim5	1	2	3	4	5	1
1	552	1204	471	0	0	**2227**
2	4	3299	3	0	0	**3306**
3	1	135	6435	0	0	**6571**
4	0	65	147	273	3	**488**
5	0	0	0	36	135	**171**
Total	557	4703	7056	309	138	12763*

Description of the five river types of environmental variability

Typology of the five river types:

1. Type: Mountainous Type:

Small sized rivers (size of catchment, wetted witdth) with high altitude and slope in mountainous areas of Europe (e.g. sites in Switzerland, Austria, Spain, France, Italy).

2. Type: Mediterranean Type:

Small sized rivers (size of catchment, wetted witdth) with medium altitude and low slope in the south-western part of Europe (e.g. Portugal, Spain and France).

3. Type: Central Type:

Medium sized rivers with medium altitude and low slope in central Europe (lower rivers, e.g. UK, Romania, Poland, Lithuania, Hungary, Scandinavia).

4. Type:

Large sized rivers with large catchment, low altitude and slope in Central/Western Europe (Netherlands, Germany, Poland).

5. Type:

Very large sized rivers with very large catchment, and extreme low slope and altitude in Central/Western Europe (Netherlands and Germany).

Table 19: Distribution of sites per country and type of environmental variability

	River types					
Country	1	2	3	4	5	Total
AT	103	0	**808**	11	0	922
CH	**491**	10	216	0	0	717
DE	0	0	**456**	211	118	785
ES	**754**	**2350**	4	21	0	3129
FI	0	0	256	56	0	312
FR	409	114	**551**	67	1	1142
HU	1	0	**157**	25	5	188
IT	**254**	32	**284**	4	0	574
LT	0	0	109	6	0	115
NL	0	0	147	7	28	182
PL	25	0	**856**	20	18	919
PT	77	**800**	2	44	0	923
RO	26	0	**229**	8	0	263
SE	4	0	**593**	8	0	605
UK	83	0	**1903**	0	1	1987
Total	2227	3306	6571	488	171	12763

The error bars with 95% confidence interval from the following figures 19 to 24 are showing the environmental characteristics for the 5 river types of environmental variability, described by mean values of altitude [m], slope [%], latitude, longitude, size of catchment [km²] and wetted width [m].

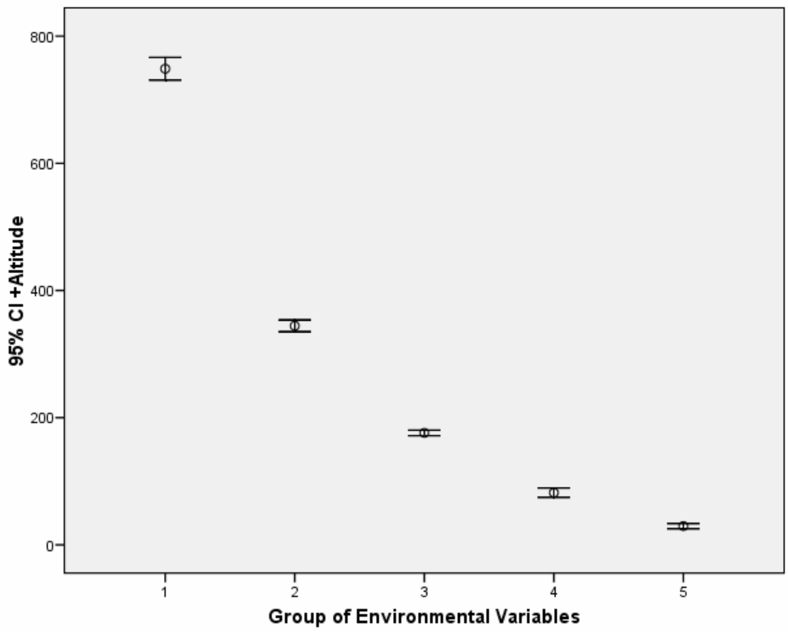

Figure 19: Error bar, altitude/group of environmental variables (error bar 95% confidential interval)

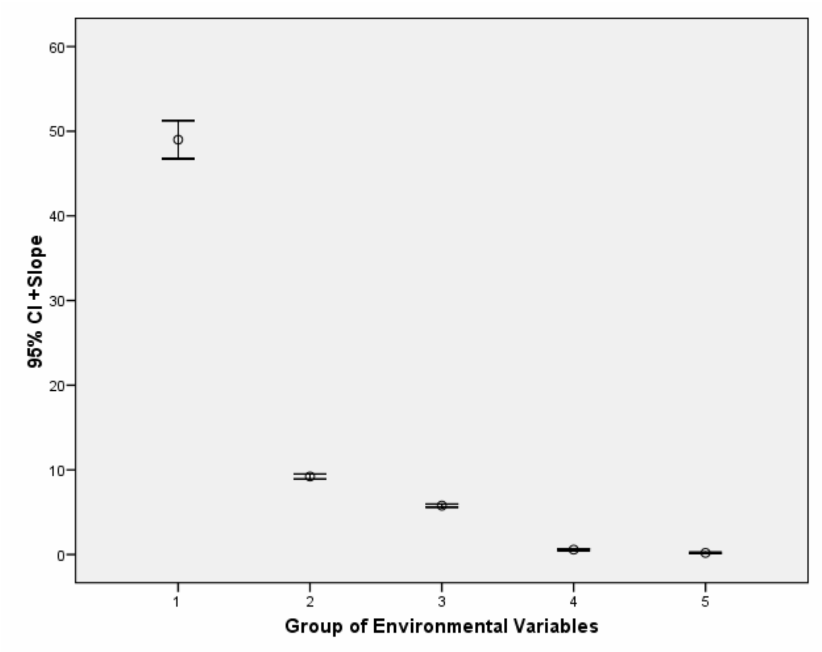

Figure 20: Error bar, slope/group of environmental variables (error bar 95% confidential interval)

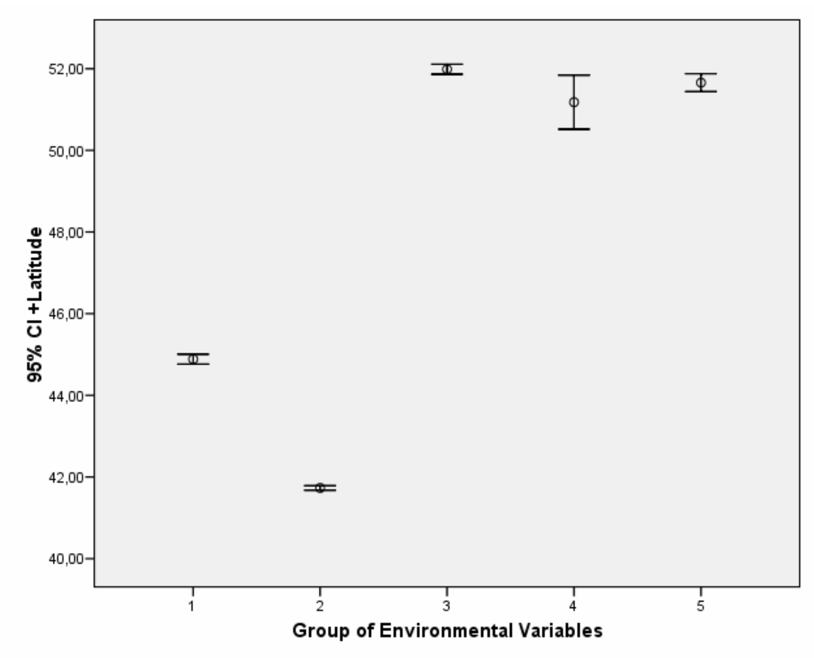

Figure 21: Error bar, latitude/group of environmental variables (error bar 95% confidential interval)

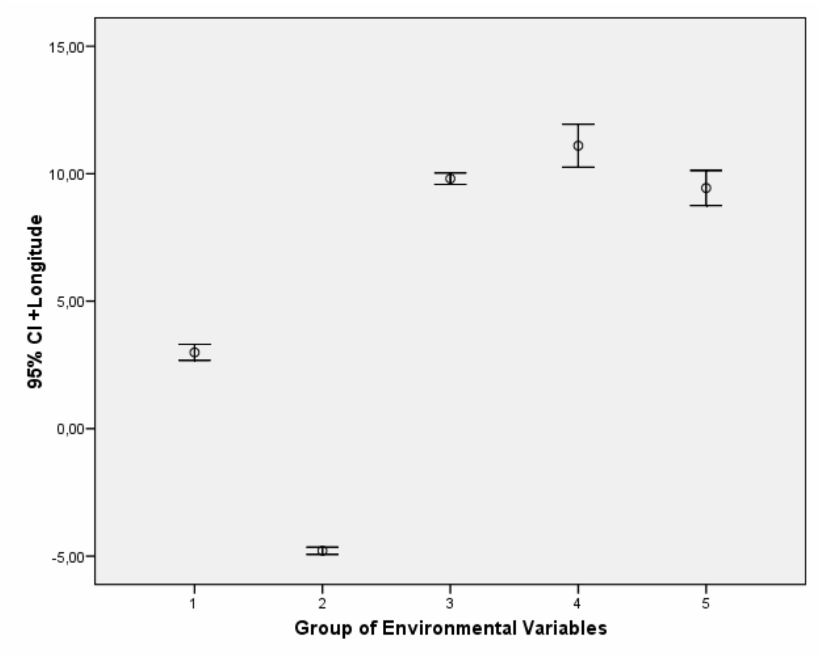

Figure 22: Error bar, longitude/group of environmental variables (error bar 95% confidential interval)

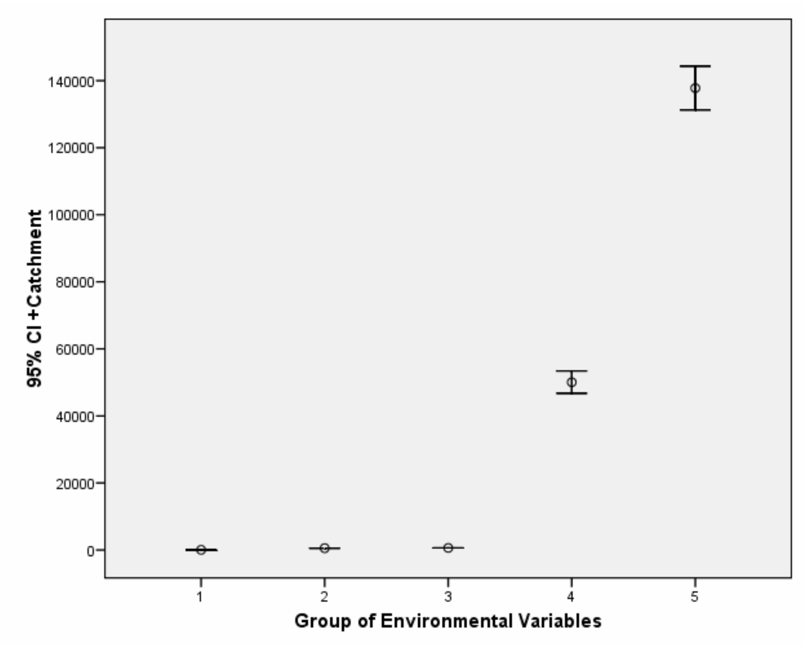

Figure 23: Error bar, size of catchment/group of environmental variables (error bar 95% confidential interval)

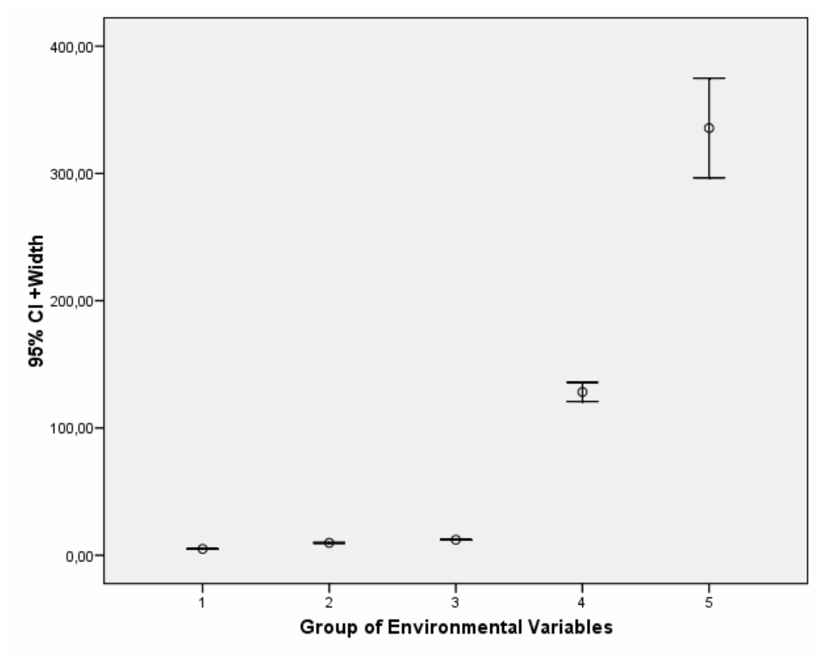

Figure 24: Error bar, wetted width/group of environmental variables (error bar 95% confidential interval)

This five river types of environmental variability were used for all further analyses of the pressure dataset. In general, the focus for index development will be set on river types 1, 2 and 3 (including 12104 sites), sites located in types 4 and 5 (including 659 sites) will be more important and interesting for assessment of large floodplain rivers. Nevertheless, the occurrence of types 4&5 must be kept in mind for all further steps of EFI+ development.

6.4.5. Variable Selection for Pressure Index Development

The PCA (principle component analysis) was used to point out redundant variables and to reduce the number of variables, which will be used for the index calculation. This PCA quantifies categorical variables while reducing the dimensionality of the data (SPSS Categories® 11.0 manual). The CATPCA (Categorical principal components analysis) analysis is also known as categorical principal components analysis.

The main goal of PCA is to reduce an original set of variables into a smaller set of uncorrelated components that represent most of the information found in the original variables. The technique is used when a large number of variables prohibit effective interpretation of the relationships between objects.

PCA was done for the following pressure groups:

1. Hydrological pressures
2. Morphological pressures

3. Water quality pressures and

4. Connectivity pressures

The 5. pressure group which gives information on collinear connected reservoir and navigation has been excluded from pressure analyses because of its minor importance at this step.

First, PCA has been done for each group with the total dataset and afterwards with a split dataset (divided for all river fives types of environmental variability). To test and specify the outputs of PCA, correlation analysis and cross tables (for variables that potentially can be summed up) have been done in the next step.

Hydrological pressures

The following six hydrolocical variables were used for the PCA analyses: 1. Impoundment (H_imp), 2. Hydropeaking (H_hydrop) 3. Water abstraction (H_waterabstr) 4. Hydrograph modification (H_hydromod) 5. Reservoir flushing (H_resflush) and 7. Temperature impact (H_tempimp). The description why the variable Velocity increase (H_veloincr) was not used can be found on the page 81. The variables are described in the chapter: 7. Description of Austrian Variables of the EFI+ Project. The PCA is also done for the scenario Consensus 1 (8636 sites) to find out if it behaves different from the dataset 13085 sites.

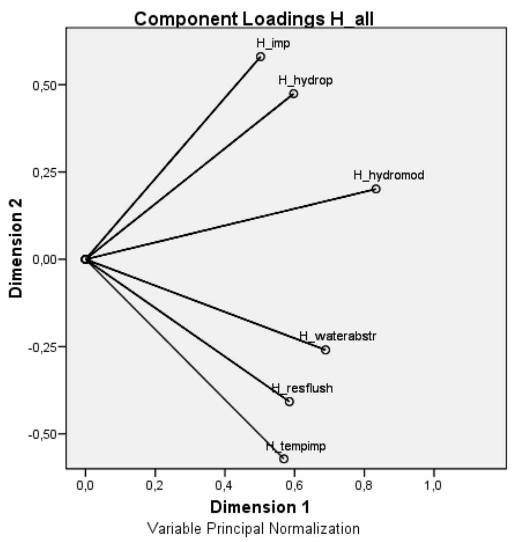

Figure 25: PCA with the variables in group "hydromorphological pressures", (dataset 13085 sites) except H_veloincr

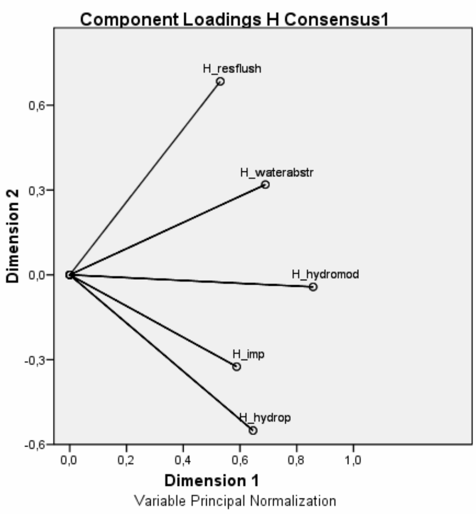

Figure 26: PCA with the variables in group "hydromorphological pressures", (dataset: Consensus 1 8636 sites) except H_veloincr

The PCA has been performed for whole group "hydrological pressures" and for the five river types of environmental variability. The conclusion of this analyses is that all hydrological variables seem to act in a different dimension/way (also in the 5 river types) Therefore they have to be integrated into the pressure type specific index and the overall index separately.

For the comparison of the PCA with the datasets 13085 sites and the dataset consensus 1 (8636 sites) no major differences where detected.

Table 20: Spearman's rank correlation for all hydrological pressures

		H_hydrop	H_waterabstr	H_hydromod	H_tempimp	H_resflush	H_imp	H_veloincr
H_hydrop	Corr_Coeff	1.000	.217(**)	.487(**)	.145(**)	.118(**)	.277(**)	-.036(**)
	Sig	_	.000	.000	.000	.000	.000	.000
	N	12821	11678	12677	10670	12778	12087	10439
H_waterabstr	Corr_Coeff	.217(**)	1.000	.427(**)	.416(**)	.274(**)	.181(**)	.248(**)
	Sig)	.000	_	.000	.000	.000	.000	.000
	N	11678	11688	11633	9606	11647	10958	9643
H_hydromod	Corr_Coeff	.487(**)	.427(**)	1.000	.313(**)	.351(**)	.366(**)	.078(**)
	Sig	.000	.000	_	.000	.000	.000	.000
	N	12677	11633	12687	10617	12645	11957	10390
H_tempimp	Corr_Coeff	.145(**)	.416(**)	.313(**)	1.000	.363(**)	.055(**)	.216(**)
	Sig	.000	.000	.000	_	.000	.000	.000
	N	10670	9606	10617	10689	10649	10670	10362
H_resflush	Corr_Coeff	.118(**)	.274(**)	.351(**)	.363(**)	1.000	.155(**)	.268(**)
	Sig	.000	.000	.000	.000	_	.000	.000
	N	12778	11647	12645	10649	12798	12055	10412
H_imp	Corr_Coeff	.277(**)	.181(**)	.366(**)	.055(**)	.155(**)	1.000	.019(*)
	Sig	.000	.000	.000	.000	.000	_	.048
	N	12087	10958	11957	10670	12055	12097	10442
H_veloincr	Corr_Coeff	-.036(**)	.248(**)	.078(**)	.216(**)	.268(**)	.019(*)	1.000
	Sig	.000	.000	.000	.000	.000	.048	_
	N	10439	9643	10390	10362	10412	10442	10454

** Correlation is significant at the 0.01 level (2-tailed).
* Correlation is significant at the 0.05 level (2-tailed).

The Spearman's rank correlation for "hydrological pressures", shows that none of them are highly correlated.

Furthermore the relationship between hydrological and morphological variables has been analysed. The result was that variable H_veloincr (velocity increase due to channelisation and floodprotection), which is a part of the hydrological pressures group, fits better to instream channel variables in the section morphology. Therefore, variable H_veloincr has been recoded into M_H_veloincr and henceforth is consitered as morphological variable.

The conclusion for hydrological pressures is that following variables are qualified for pressure index calculation:

1. H_imp
2. H_hydrop
3. H_waterabstr
4. H_hydromod
5. H_resflush and
6. H_tempimp

Morphological pressures

The following eight morphological variables were used for the PCA analyses of morphological pressures: 1. Channelisation (M_channel), 2. Cross section (M_crosssec) 3. Instream habitat (M_instrhab), 4. Hydrograph modification (H_hydromod), 5. Riparian vegetation (M_ripveg), 6. Floodprotection (M_floodpr), 7. Sedimentation (M_sediment) and 8. Remaining floodplain (M_remfloodpl). The variables are described in the chapter: 7. Description of Austrian Variables.

The PCA for this pressure group has been done for the whole group and also for the five types of natural variability. As mentioned before, the variable velocity increase is included into the following analyses.

The figures 27 and 29 show that M_sed and M_remfloodpl (orange arrows) represent different dimensions. The figure 30 shows that some variables might be redundant (orange ellipse).

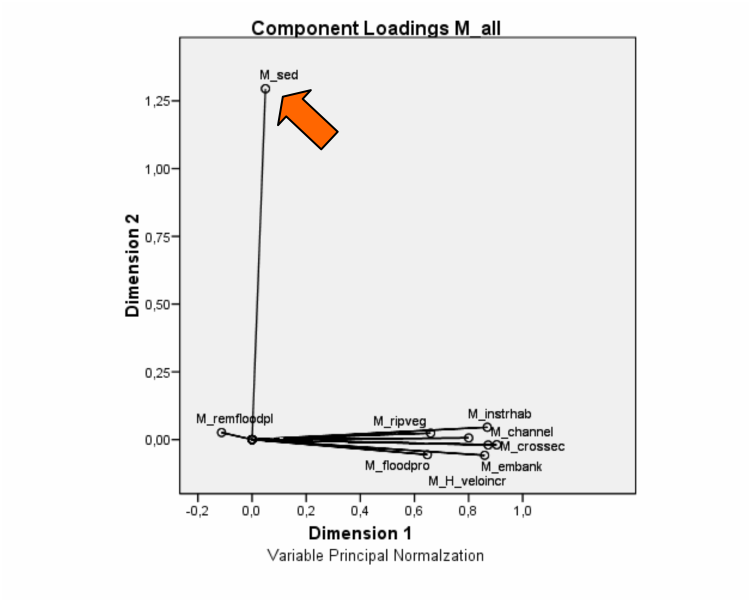

Figure 27: PCA with the variables in the group "morphological pressures" (dataset: 13085 sites)

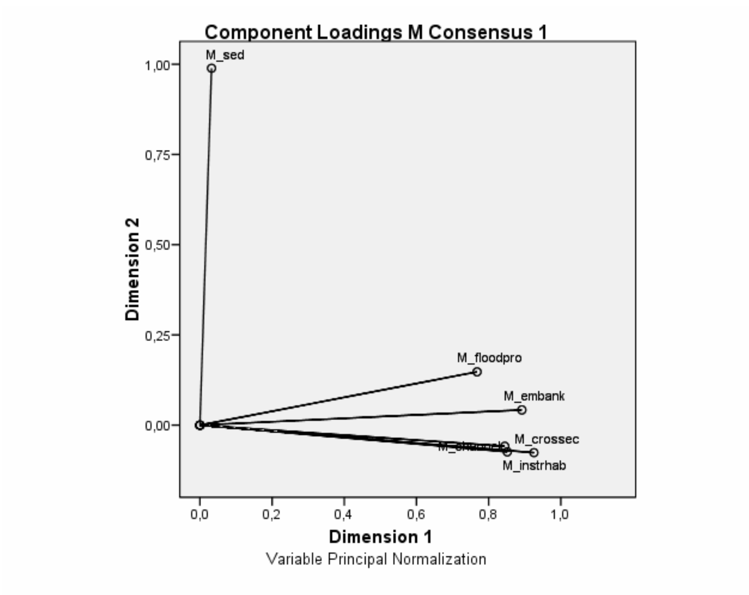

Figure 28: PCA with the variables in the group "morphological pressures" (dataset: Consensus 1 8636 sites)

The comparison of the PCA with the datasets 13085 sites and the dataset consensus 1 shows no major differences in-between the correlations of the variables.

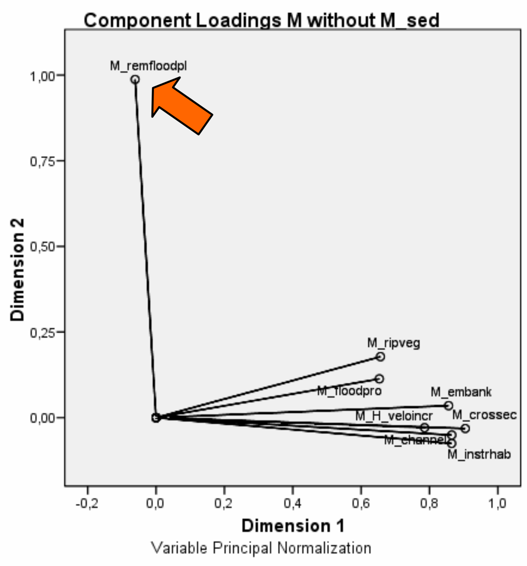

Figure 29: PCA with the variables in the group “morphological pressures”, without M_sed (dataset: 13086 sites)

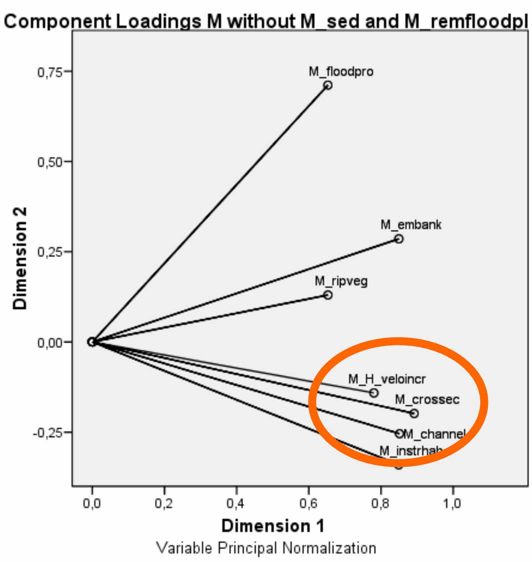

Figure 30: PCA with the variables in the group “morphological pressures”, without M_sed and M_remfloodpl (dataset: 13086 sites)

The first gained recognition of PCA was that the variables M_channel, M_instrhab, M_H_veloincr and M_crossec go into the same direction and therefore represent the same dimension. To get clear correlations between these variables, M_sed and M_remfloodpl have been excluded from PCA, because of acting in a different dimension.

In order to test the hypothesis that M_channel, M_instrhab, M_H_veloincr and M_crossec are highly correlated, analyses with cross tables and correlation analysis have been done.

Different to the redundant variables, M_embankm, M_ripveg, M_sediment, M_floodpr and M_remfloodpl act in a separate dimension/way and will be used separately for index calculation. Eccept M_remfloodpl, they are no more treated by further analysis in this section. Correlation analyses and cross tables for identifying redundant morphological variables are given in the table 21.

Table 21: Spearman's rank correlation for the four variables M_channel, M_instrhab, M_H_veloincr and M_crossec, showing that they are highly correlated.

		M_channel	M_instrhab	M_H_veloincr	M_crossec
M_channel	Correlation Coefficient	1.000	**.633**	**.704**	**.710**
	Sig_ (2-tailed)		.000	.000	.000
	N	12096	11900	10393	11930
M_instrhab	Correlation Coefficient	**.633**	1.000	**.623**	**.728**
	Sig_ (2-tailed)	.000		.000	.000
	N	11900	11943	10201	11913
M_H_veloincr	Correlation Coefficient	**.704**	**.623**	1.000	**.595**
	Sig_ (2-tailed)	.000	.000		.000
	N	10393	10201	10454	10228
M_crossec	Correlation Coefficient	**.710**	**.728**	**.595**	1.000
	Sig_ (2-tailed)	.000	.000	.000	
	N	11930	11913	10228	11973

.

According to the previous table 21, M_channel, M_instrhab, M_H_veloincr and M_crossec can be summarized into one variable because they are highly correlated. This decision is also supported by the following cross tables. The number of cases in crosstables should always be at least 70% of total values. To test variables with 2 categories against variables with 3 categories, last ones have been summed up into 2 categories.

Table 22: Cross table relationship between variable velocity increase and channelisation (summed up in 2 categories for comparison; 1, 3 =1, 5 =5).

		M_H_veloincr		Total
M_channel_sumup2		1 No	3 Yes	1 No
1	Count	8017	820	8837
	% within M_channel_sumup2	**90.7%**	9.3%	100.0%
5	Count	400	1156	1556
	% within M_channel_sumup2	25.72%	**74.3%**	100.0%
Total	Count	8417	1976	10393
	% within M_channel_sumup2	81.0%	19.0%	100.0%

The consistency between the variable M_channel_sumup2 and M_H_veloincr is in both cases more than 70%.

Table 23: Cross table relationship between variable velocity increase and alteration of cross-section (summed up in 2 categories for comparison; 1, 3 = 1, 5 =5).

		M_H_veloincr		Total
M_crossec_sumup2		1 No	3 Yes	1 No
1	Count	7882	808	7074
	% within M_crossec_sumup2	**90.7%**	9.3%	100.0%
5	Count	410	1128	3154
	% within M_crossec_sumup2	26.7%	**73.3%**	100.0%
Total	Count	8292	1936	10228*
	% within M_crossec_sumup2	81.1%	18.9%	100.0%

The consistency between variable velocity increase and alteration of cross-section is in both cases more than 70%.

Table 24: Cross table relationship between variable instream habitat and alteration of cross-section.

		M_instrhab			Total
M_crossec		1 No	3 Intermediate	5 High	1 No
1 No	Count	6499	1429	174	8102
	% within M_crossec *M_crossec	**80.2%**	17.6%	2.1%	100.0%
3 Intermediate	Count	346	1379	395	2120
	% within M_crossec *M_crossec	16.3%	65.0%	18.6%	100.0%
5 Technical crosssec	Count	59	240	1392	1691
	% within M_crossec *M_crossec	3.5%	14.2%	**82.3%**	100.0%
Total	Count	6904	3048	1961	11913
	% within M_crossec *M_crossec	58.0%	25.6%	16.5%	100.0%

The consistency between the variable instream habitat and alteration of cross-section is in both cases more than 70%.

Table 25: Cross table relationship between variable instream habitat and channelisation.

		M_instrhab			Total
M_channel		1 No	3 Intermediate	5 High	1 No
1 No	Count	6387	1695	360	8442
	% within M_channel	**75.7%**	20.1%	4.3%	100.0%
3 Intermediate	Count	447	884	415	1746
	% within M_channel	25.6%	50.6%	23.8%	100.0%
5 Straightened	Count	69	440	1203	1712
	% within M_channel	4.0%	25.7%	**70.3%**	100.0%
Total	Count	6903	3019	1978	11900
	% within M_channel	58.0%	25.4%	16.6%	100.0%

The consistency between variable instream habitat and channelisation is in both cases more than 70%.

Variables describing floodplains:

The variable M_remfloodpl always creates a different PCA dimension. The difference to other variables in morphological variable group is that it gives a diametric information (high = best case). This variable has been compared with two other “floodplain” variables of the dataset: M_floodpr (floodprotection) and floodplain_site (presence of former floodplain, Table site) to find out if their outputs fit together. Table 26 shows that especially M_remfloodpl and floodplain_site do not fit together. 68.8 % of their cases for M_remfloodpl are not well classified, when compared with the information about former floodplain (category No in M_remfloodpl means “no more floodplain available”, NoData would indicate that no former floodplain ever existed. – This was agreed in the consortium during data correction in November 2007.

Table 26: Cross table relationship between M_remfloodpl and Floodplain_site.

		Floodplain						Total
M_remfloodpl		Large	Medium	Small	Some water	No	NoData	
No	Count	7	11	32	12	**5762**	2556	8380
	% within Floodplain_site	.1%	.1%	.4%	.1%	**68.8%**	30.5%	100.0%
Yes	Count	365	435	682	70	1277	812	3641
	% within Floodplain_site	10.0%	11.9%	18.7%	1.9%	35.1%	22.3%	100.0%
Total	Count	392	449	715	89	7127	4313	13085*
	% within Floodplain_site	3.0%	3.4%	5.5%	.7%	54.5%	33.0%	100.0%

Due to these facts and the “low” completeness of 75% of the variable M_remfloodpl (available only for of the pressure analysis dataset), it was proposed not to use this variable for the index development.

Conclusion for morphological pressures:

A new variable (M_floodpl_total) for further floodplain analyses will be created: In cases where variable flooplain_site is filled with “no”, this information will be adopted, in all other cases information out of variable M_remfloodplain will be taken.

Redundant variables M_channel, M_instrhab, M_H_veloincr and M_crossec will be summed up in variable M_morph_instr (by taking the Mean of these 4 variables).

Variables qualified for pressure index calculation:

1. Morph_instream,
2. M_embankm
3. M_ripveg
4. M_floodpr and
5. M_sediment

Water quality Pressures

The following six water quality pressure variables were used for the PCA analyses of morphological pressures: 1. acidification (W_acid), 2. water quality index (W_index), 3. toxic substances (W_toxic), 4. organic siltation (W_osilt), 5. organic pollution (W_opoll) and 6. eutrophication (W_eutroph) The variables are described in the chapter: 7. Description of Austrian Variables

The PCA analyses, for all variables out of “water quality pressures” show that W_index and W_acid act in different dimension (orange arrows), W_osilt and W_toxic as well as W_eutrph and W_poll seem to be redundant.

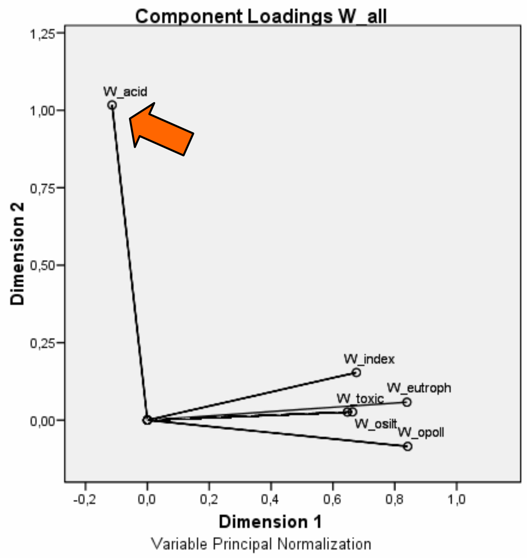

Figure 31: PCA with the variables in group “water quality pressures” (13086 sites)

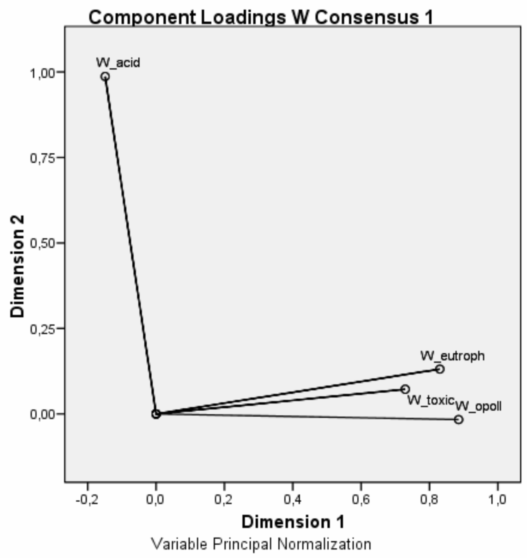

Figure 32: PCA with the variables in group “water quality pressures”, (dataset: Consensus 1, 8636 sites)

For the comparison of the PCA with the datasets 13085 sites and the dataset consensus 1 no major differences can be detected.

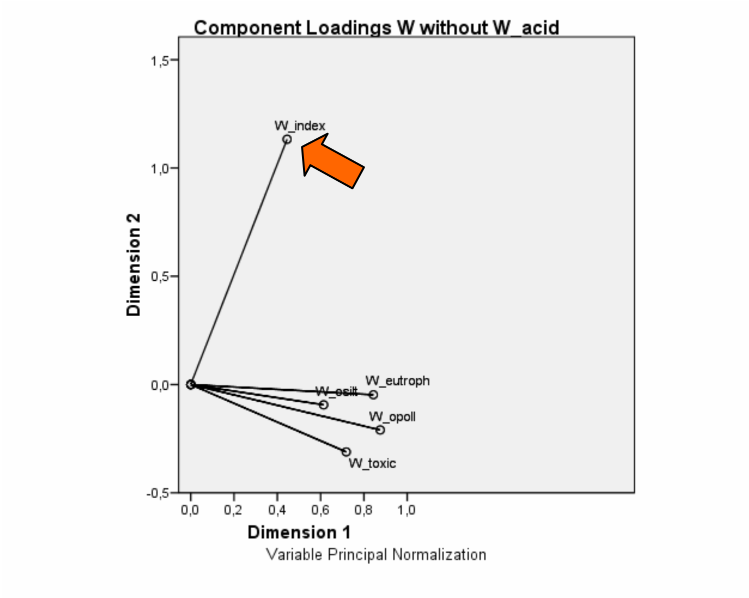

Figure 33: PCA with the variables in group “water quality pressures” without W-acid (13086 sites)

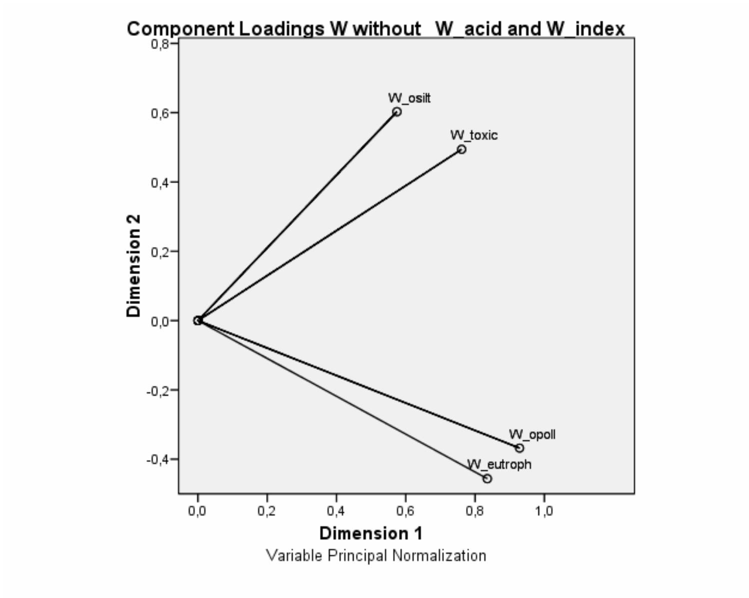

Figure 34: PCA with the variables in group “water quality pressures” without W-acid and W-Index (13086 sites)

The table 27 tests the hypothesis of the PCA with a Spearman’s rank correlation. Some variables show slight correlations. Due to the fact that they describe totally different effects, all pressure variables of the water quality are integrated separately into the pressure index.

Table 27: Spearman's rank correlations between all variables of group "water quality pressures".

		W_toxic	W_acid	W_index	W_eutroph	W_opoll	W_osilt
W_toxic	Correlation Coefficient	1.000	.008	.166(**)	.327(**)	.418(**)	**.523(**)**
	Sig_ (2-tailed)		.378	.000	.000	.000	.000
	N	11653	11637	10899	11599	11433	9588
W_acid	Correlation Coefficient	.008	1.000	-.020(*)	-.030(**)	-.099(**)	-.057(**)
	Sig_ (2-tailed)	.378		.031	.001	.000	.000
	N	11637	12768	11755	12028	11927	10105
W_index	Correlation Coefficient	.166(**)	-.020(*)	1.000	**.529(**)**	.423(**)	.410(**)
	Sig_ (2-tailed)	.000	.031		.000	.000	.000
	N	10899	11755	11783	11135	11245	9259
W_eutroph	Correlation Coefficient	.327(**)	-.030(**)	**.529(**)**	1.000	**.603(**)**	.425(**)
	Sig_ (2-tailed)	.000	.001	.000		.000	.00
	N	11599	12028	11135	12094	11783	9983
W_opoll	Correlation Coefficient	.418(**)	-.099(**)	.423(**)	**.603(**)**	1.000	.410(**)
	Sig_ (2-tailed)	.000	.000	.000	.000		.000
	N	11433	11927	11245	11783	11962	9887
W_osilt	Correlation Coefficient	**.523(**)**	-.057(**)	.410(**)	.425(**)	.410(**)	1.000
	Sig_ (2-tailed)	.000	.000	.000	.000	.000	
	N	9588	10105	9259	9983	9887	10175*

** Correlation is significant at the 0.01 level (2-tailed).

* Correlation is significant at the 0.05 level (2-tailed).

<u>The Special Case: Water Quality Index</u>

The variable W_index (=Waterquality index in 5 classes) is a more or less problematic value. The reason for this is because of different indices applied in partner-countries (chemic and biotic) have been brought together. The following cross tables compare this variable with some other important water quality-variables to get an idea if the water quality index is able to detect other water quality problems.

Table 28:Relation between water quality index and a combined variable of eutrophication and organic pollution

		W_eutoph_opoll_max				Total
Waterquality Index		1	3	4	5	1
1	Count	3009	**122**	**26**	**110**	3267
	% within W_index *W_index	92.1%	**3.7%**	**.8%**	**3.4%**	100.0%
2	Count	1753	**1508**	**195**	**357**	3813
	% within W_index *W_index	46.0%	**39.5%**	**5.1%**	**9.4%**	100.0%
3	Count	683	1378	391	284	2736
	% within W_index *W_index	25.0%	50.4%	14.3%	10.4%	100.0%
4	Count	171	369	238	183	961
	% within W_index *W_index	17.8%	38.4%	24.8%	19.0%	100.0%
5	Count	74	166	141	143	524
	% within W_index *W_index	14.1%	31.7%	26.9%	27.3%	100.0%
Total	Count	5690	3543	991	1077	11301
		50.3%	31.4%	8.8%	9.5%	100.0%

The table 28 shows that W_index is not able to detect eutroph_opll problems correctly, 10% of waterquality class 1 and 55% of waterquality class 2 are impacted by eutrophication and/or organic pollution (at least 3 = weak).

Table 29: Relation between water quality index and acidification.

		Acidification			Total
Water quality Index		No	Yes	Missing	No
1	Count		99	7	3274
	% within Water_quality_index		3.0%	.2%	100.0%
2	Count		115	11	3982
	% within Water_quality_index		2.9%	.3%	100.0%
3	Count	2812	50	6	2868
	% within Water_quality_index	98.0%	1.7%	.2%	100.0%
4	Count	1032	25	3	1060
	% within Water_quality_index	97.4%	2.4%	.3%	100.0%
5	Count	578	20	1	599
	% within Water_quality_index	96.5%	3.3%	.2%	100.0%
Total	Count	11446	309	28	11783
	% within Water_quality_index	97.1%	2.6%	.2%	100.0%

Table 30: Relation between water quality index and toxic substances.

		Toxic_substances				Total
Waterquality Index		No	Intermediate	High concentration	Missing	No
1	Count	2366	**498**	**235**	175	3274
	% within Water_quality_index	72.3%	**15.2%**	**7.2%**	5.3%	100.0%
2	Count	2769	**705**	**296**	212	3982
	% within Water_quality_index	69.5%	**17.7%**	**7.4%**	5.3%	100.0%
3	Count	1315	**808**	**457**	288	2868
	% within Water_quality_index	45.9%	28.2%	15.9%	10.0%	100.0%
4	Count	651	214	66	129	1060
	% within Water_quality_index	61.4%	20.2%	6.2%	12.2%	100.0%
5	Count	284	151	84	80	599
	% within Water_quality_index	47.4%	25.2%	14.0%	13.4%	100.0%
Total	Count	7385	2376	1138	884	11783*
	% within Water_quality_index	62.7%	20.2%	9.7%	7.5%	100.0%

The table 30 shows that W_index is not able to detect problems with toxics substances correctly, 20% of water quality class 1 and class 2 are intermediate of highly impacted by toxic substances.

Conclusion for waterquality pressures:

Because water quality index is not able to detect other important water quality pressures correctly the variable is not used for pressure index calculation,

After testing the outputs of PCA with a correlation analysis, it has been decided to integrate all remaining water quality variables separately into the pressures index. Even when some variables show correlations, they describe different effects.

Variables qualified for pressure index calculation:

1. W_toxic
2. W_acid
3. W_osilt
4. W_opol
5. W_eutroph

Connectivity Pressures

The following 7 morphological variables were used for the PCA analyses of Connectivity pressures: 1. barriers on catchment down (C_catch_down) 2. barriers on segment down (C_B_s_do) 3. barriers on segment up (C_B_s_up) 4. number of barriers on segment down (C_Bn_s_do), 5. number of barriers on segment up (C_Bn_s_up), 6. distance to next barrier on segment down (C_Bd_s_do) and 7. distance to next barrier on segment up (C_Bd_seg_up). The variables are described in the chapter: 7. Description of Austrian Variables

Because of the very low completeness of variables describing the distance to barriers on the segment level (only 31%) and the completeness of the variables C_Bd_s_up and C_Bd_s_do, (34%) and also because of the fact that of different units used (meters and kilometres) these variables will not been taken for the index calculation.

The following PCA´s show that the variables on segment up and on segment down represent have the same dimension. The variable barriers catchment down (arrange arrow) acts in a differently dimension.

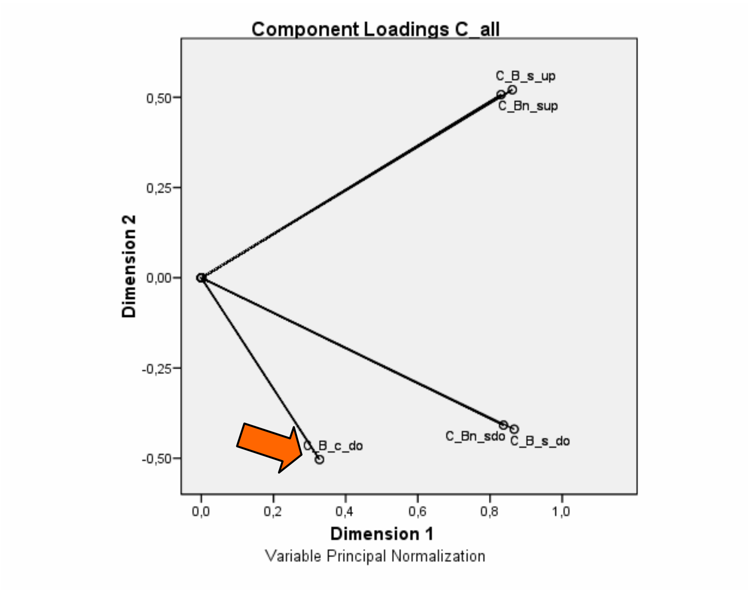

Figure 35: PCA with the variables in the group “connectivity pressures” (dataset: 13086 sites)

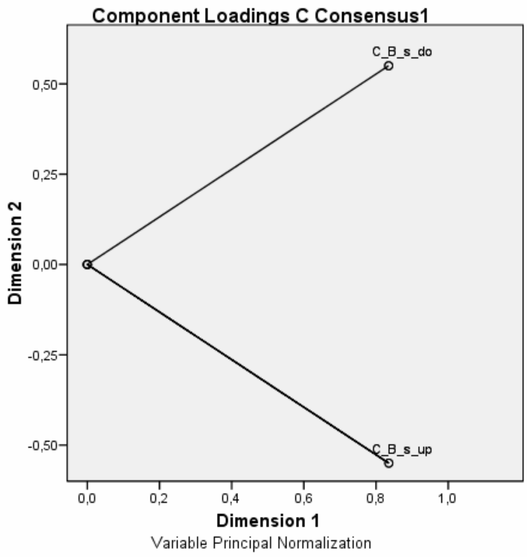

Figure 36: PCA with the variables in group “connectivity pressures” (dataset: Consensus 1 8636 sites)

For the comparison of the PCA with the datasets 13085 sites and the dataset consensus 1 no major differences can be seen.

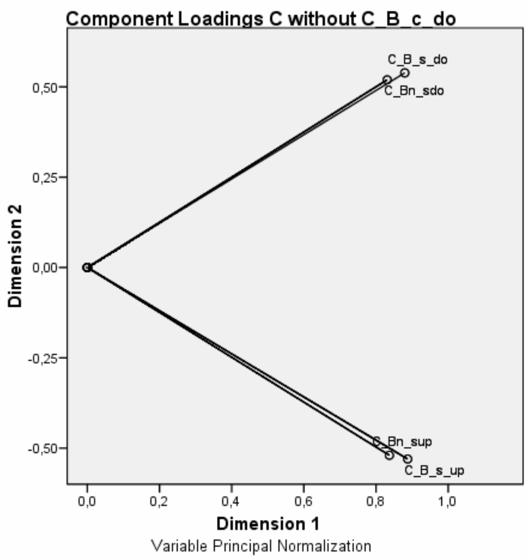

Figure 37: PCA of the variables in group “connectivity pressures”, without barriers catchment down (13086 sites)

The following table31 tests this hypothesis with Spearman’s rank correlation.

Table 31: Spearman's rank correlation between connectivity variables.

		C_B_c_do	C_B_s_up	C_B_s_do	C_Bn_sup	C_Bn_sdo
C_B_c_do	Correlation Coefficient	1.000	.139(**)	.188(**)	.183(**)	.234(**)
	Sig_ (2-tailed)		.000	.000	.000	.000
	N	12821	12750	12741	11615	11522
C_B_s_up	Correlation Coefficient	.139(**)	1.000	.466(**)	**.975(**)**	.474(**)
	Sig_ (2-tailed)	.000		.000	.000	.000
	N	12750	12764	12740	11613	11516
C_B_s_do	Correlation Coefficient	.188(**)	.466(**)	1.000	.468(**)	**.971(**)**
	Sig_ (2-tailed)	.000	.000		.000	.000
	N	12741	12740	12755	11599	11522
C_Bn_sup	Correlation Coefficient	.183(**)	**.975(**)**	.468(**)	1.000	.541(**)
	Sig_ (2-tailed)	.000	.000	.000		.000
	N	11615	11613	11599	11619	11339
C_Bn_sdo	Correlation Coefficient	.234(**)	.474(**)	**.971(**)**	.541(**)	1.000
	Sig_ (2-tailed)	.000	.000	.000	.000	
	N	11522	11516	11522	11339	11525

** Correlation is significant at the 0.01 level (2-tailed).

The PCA analyses show that both variables describing the segment down and both variables describing the segment are highly correlated.

Conclusion for connectivity pressures:

Two new variables C_seg_up (Mean of C_bn_seg_up and C_b_seg_up) as well as C_seg_down (Mean of C_bn_seg_down and C_b_seg_down) will be created.

Variable C_catch_down instead creates a different dimension and should not be consitered for index development. The reason is its large spatial extension (indicates if there is a barrier downstream on unit catchment until sea).

Variables qualified for pressure index calculation:

1. C_seg_down
2. C_seg_up.

The Table 32 describes the retaining variables for the variables for the index development for all four-pressure types. Additionally in the field under each variable and category the frequencies of occurrence in % are indicated.

Table 32: Retaining variables for the index development

Pressure Group	Variable	Abbrevation of Variable	Intensity				
Hydrology	Impoundment	H_imp	no (1)	weak (3)	stong (5)		
		Freqency in %	88.5	6.2	5.3		
	Hydropeaking	H_hydrop	no (1)	partial (3)	yes (3)		
		Freqency in %	92.3	7.7			
	Waterabstraction	H waterabstr.	no (1)	weak to medium (3)	stong (5)		
		Freqency in %	71.0	17.2	11.8		
	Reservoir Flushing	H_resflush	no (1)	yes (3)			
		Freqency in %	96.1	3.9			
	Hydromodification	H_hydromod	no (1)	yes (3)			
		Freqency in %	81.8	18.2			
	Temperature impact	H_tempimp	no (1)	permanent increase (3)	permanent decrease (3)	summer increase (3)	summer decrease (3)
		Freqency in %	89.5	10.5			
Morphology	Morphology Instream	M_channel	no (1)	intermediate (3)	straightened (5)		
		Freqency in %	71.1	14.8	14.2		
		M_crossec	no (1)	intermediate (3)	technical crossec (5)		
		Freqency in %	67.9	18.0	14.1		
		H_veloincr	no (1)	yes (3)			
		Freqency in %	81.1	18.9			
		M_instrhab	no (1)	intermediate (3)	high (5)		
		Freqency in %	57.8	25.5	16.6		
	Embankment	M_embankm	no (1)	slight (2)	intermediate (3)	high (5)	
		Freqency in %	65.8	13.1	10.9	10.1	
	Riparian Vegetation	M_ripveg	no (1)	slight (2)	intermediate (3)	high (5)	
		Freqency in %	45.0	17.0	19.8	18.2	
	Floodprotection	M_floodpro	no (1)	yes (3)			
		Freqency in %	77.0	23.0			
	Sedimentation	M_sed	no (1)	weak (2)	medium (3)	high (5)	
		Freqency in %	56.7	25.5	12.0	5.8	
Water Qualitay	Organic Siltation	W_orgsilt	no (1)	yes (3)			
		Freqency in %	80.9	19.1			
	Toxic Substances	W_toxic	no (1)	weak (3)	high (5)		
		Freqency in %	68.3	21.6	10.1		
	Acidification	W_acid	no (1)	yes (3)			
		Freqency in %	97.5	2.5			
	Eutrophication	W_eutroph	no (1)	low(3)	intermediate (4)	extreme (5)	
		Freqency in %	59.0	22.4	16.9	1.7	
	Organic Pollution	W_opoll	no (1)	weak (3)	strong (5)		
		Freqency in %	62.8	28.6	8.6		
Connectivity	Conectivity Segment up	C_bs_up (Barrier Segment up)	no (1)	partial (3)	yes (3)		
		Freqency in %	76.8	23.2			
		C_nbs_up (Barrier Number Segment up)	no (1)	medium (3)	high (4)		
		Freqency in %	76.4	17.6	5.3		
	Conectivity Segment Down	C_bs_down (Barrier Segment down)	no (1)	partial (3)	yes (3)		
		Freqency in %	78.0	22.0			
		C_nbs_down (Barrier Number Segment down)	no (1)	medium (3)	high (4)		
		Freqency in %	77.7	16.4	5.2		

7. Pressure index development

In the previous chapter chosen variables for the index calculation are combined to pressure type indices (e.g. hydrology, morphology, water quality, continuity) and later on aggregated into a global pressure index.

The table 33 gives an overview of the different index scenarios and the used variables and the index scenarios are described in the chapter: 6.4.2 Index Scenarios.

Table 33: Scenarios for index calculation

Scenarios	Max Data Loss	Consensus 1	Consensus 2	Min Data Loss
Country	No. of sites	No. of sites	No. of sites	No. of sites
AT	847	873	879	879
CH	**No data left**	**No data left**	70	428
DE	741	741	764	772
ES	1650	1596	1596	1613
FI	257	257	257	257
FR	**No data left**	594	874	808
HU	166	166	166	166
IT	303	315	468	469
LT	54	83	83	90
NL	182	182	182	182
PL	884	888	893	893
PT	923	923	923	923
RO	262	262	262	262
SE	602	602	602	602
UK	**No data left**	1154	1154	1154
No.Sites	**6871**	**8636**	**9173**	**9498**
	Variables kept for the Index Calculation			
1	Fished_area	Fished_area	Fished_area	Fished_area
2	Wetted_width	Wetted_width	Wetted_width	Wetted_width
3	H_imp	H_imp	H_imp	H_imp
4	H_hydrop	H_hydrop	H_hydrop	H_hydrop
5	H_waterabstr	H_waterabstr	H_waterabstr	H_waterabstr
6	H_hydromod	H_hydromod	H_hydromod	H_hydromod
7	H_tempimp	H_resflush	H_resflush	H_resflush
8	H_veloincr	M_sed	M_sed	M_channel
9	H_resflush	M_channel	M_channel	M_crossec
10	M_sed	M_crossec	M_crossec	M_instrhab
11	M_channel	M_instrhab	M_instrhab	M_embank
12	M_crossec	M_embank	M_embank	M_floodpro
13	M_instrhab	M_floodpro	M_floodpro	C_B_s_up
14	M_ripveg	C_B_s_up	C_B_s_up	C_B_s_down
15	M_embank	C_B_s_down	C_B_s_down	W_acid
16	M_floodpro	W_toxic	Distance_from_source	Distance_from_source
17	W_toxic	W_acid	Size_of_catchment	Size_of_catchment
18	W_acid	W_eutroph	Altitude	Altitude
19	W_eutroph	W_opoll	Actual_river_slope	Actual_river_slope
20	W_opoll	Distance_from_source	Geomorph_river_type	Geomorph_river_type
21	W_osilt	Size_of_catchment	Flow_regime	Flow_regime
22	C_B_s_up	Altitude	Floodplain	Floodplain
23	C_B_s_do	Actual_river_slope	Valley_form	Valley_form
24	Distance_from_source	Geomorph_river_type	Natural_sediment	Natural_sediment
25	Size_of_catchment	Flow_regime	Lakes upstream	Lakes upstream
26	Altitude	Floodplain	Run1	Run1
27	Actual_river_slope	Valley_form		
28	Valley_form	Natural_sediment		
29	Geomorph_river_type	Lakes upstream		
30	Flow_regime	Run1		
31	Watersource type			
32	Floodplain			
33	Valley_slope			
34	Run1			

Multiple Impact of the Index Variants

For the comparison and to get an overview about the intensity and number of impacts per site, the following analyses have been performed for all five Index scenarios and the three index variants. The Index scenarios and variants are described in the chapter: 6.4.2 Index Scenarios.

The figures 38, 40, 42, 44, and 46 describe the mean number of pressures per country, index variant and index scenarios.

The figures 39, 41, 43, 45 and 47 describe the amount of pressures per country for each index variant and index scenarios.

In the tables, 34, 36, 37, 38 and 40 the missing, mean und median values for all for each index variant and index scenarios are described.

In the tables, 35,37,39,41 and 42 the frequency and percentage for all for each index variant and index scenarios are described.

1. Variant: 13085 Sites

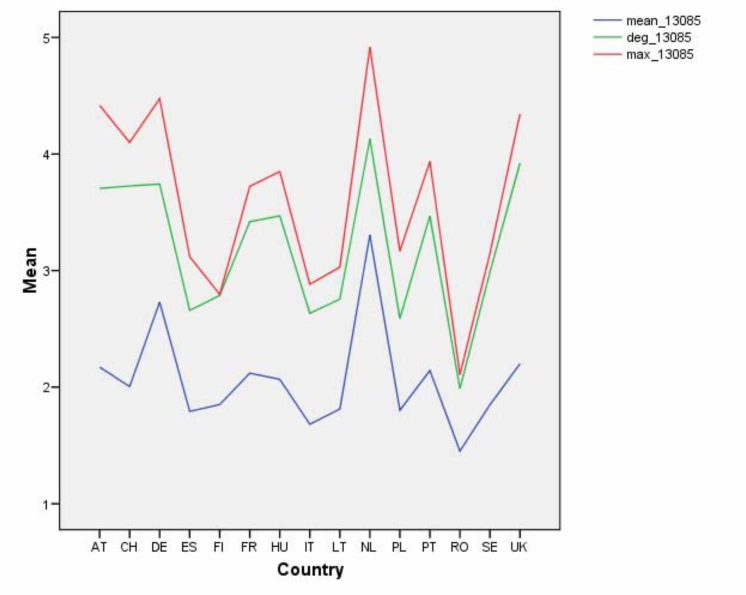

Figure 38: Distribution of the three different indices per country for the dataset with 13085 sites.

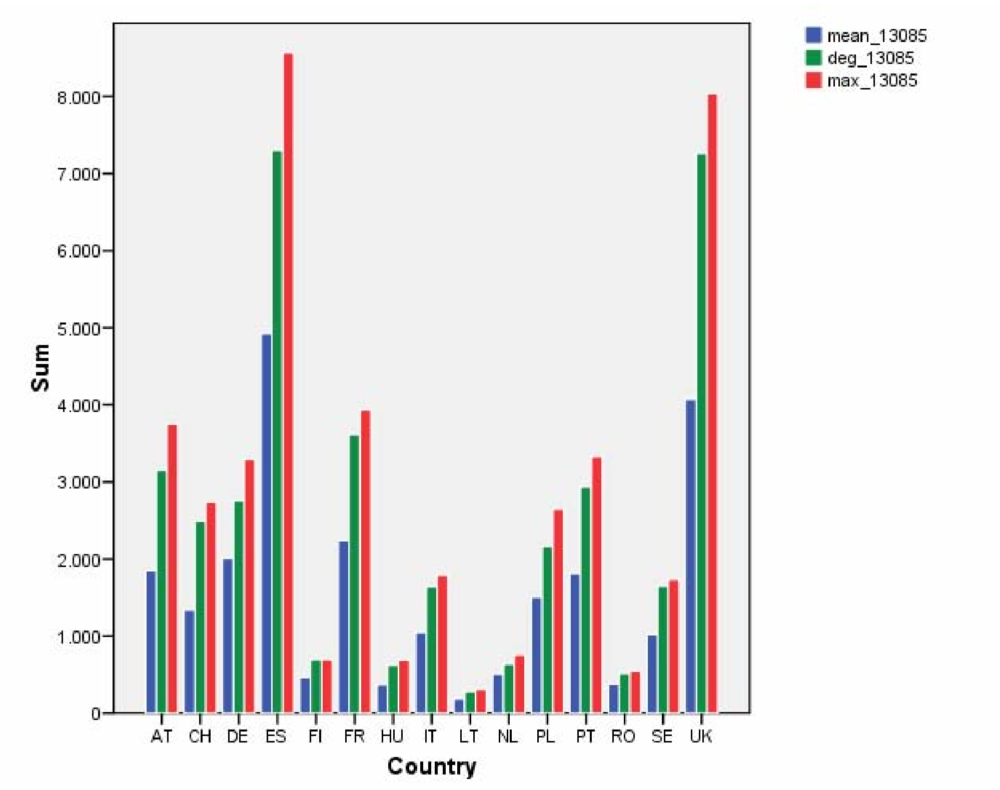

Figure 39: Summary of the 3 variants of indexes: arithmetic mean, degraded approach and worst case for the dataset 13085 sites

Table 34: Valid, missing, mean und median values for all 3 variants of indexes for the dataset 13085 sites

Statistics	Mean_13085	Deg_13085	Max_13085
Valid	11978	12310	12865
Missing	1107	775	220
Mean	2.0	3.2	3.6
Median	2.0	4.0	4.0

Table 35: Frequency and percentage for the dataset 13085 sites

	Mean_13085		Deg_13085		Max_13085	
Status	Frequency	%	Frequency	%	Frequency	%
1	1728	13.2	2257	17.2	1728	13.2
2	8264	63.2	54	0.4	589	4.5
3	1918	14.7	3507	26.8	3574	27.3
4	68	0.5	5741	43.9	1826	14.0
5			751	5.7	5148	39.3
Missing	1107	8.5	775	5.9	220	1.7
Total	13085	100	13085	100	13085	100

2. Variant, "Maximum Data Loss" (6871 Sites)

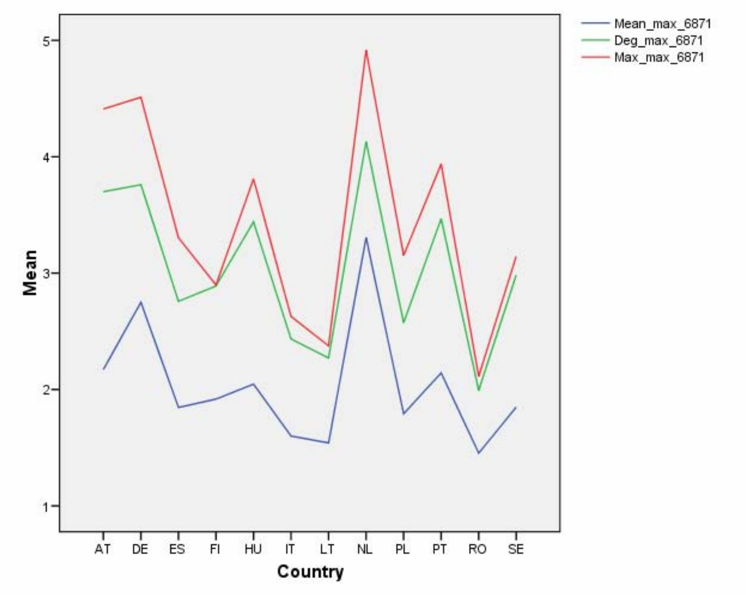

Figure 40: Distribution of the three different indices per country for the dataset maximum data loss (6871 sites).

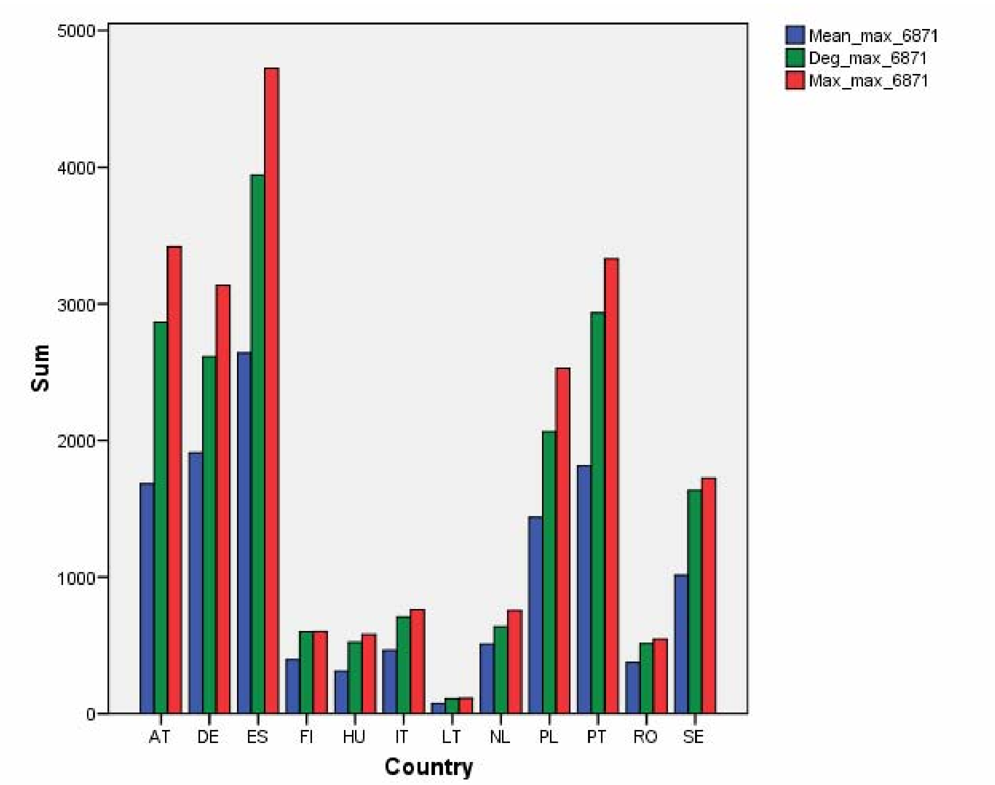

Figure 41: Summary of the 3 variants of indexes: arithmetic mean, degraded approach and worst case for the dataset maximum data loss

Table 36: Valid, missing, mean und median values for all 3 variants of indexes for the dataset maximum data loss

Statistics	**mean_max_6871**	**deg_max_6871**	**max_max_6871**
Valid	6367	6538	6871
Missing	504	333	0
Mean	2.0	3.1	3.5
Median	2.0	3.0	4.0

Table 37: Frequency and percentage for the dataset maximum data loss

	mean_max_6871		**deg_max_6871**		**max_max_6871**	
Status	Frequency	%	Frequency	%	Frequency	%
1	989	14.4	1379	20.1	989	14.4
2	4231	61.6	23	0.3	370	5.4
3	1083	15.8	2007	29.2	2023	29.4
4	64	0.9	3001	43.7	903	13.1
5			128	1.9	2586	37.6
Missing	504	7.3	333	4.8		
Total	6871	100	6871	100	6871	100

3. Variant, "Consensus 1" (8636 Sites)

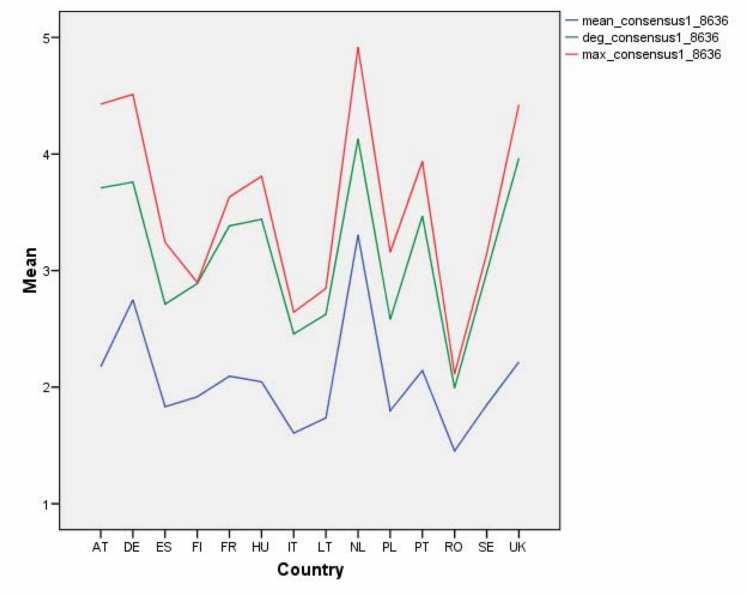

Figure 42: Distribution of the three different indices per country for the dataset consensus 1.

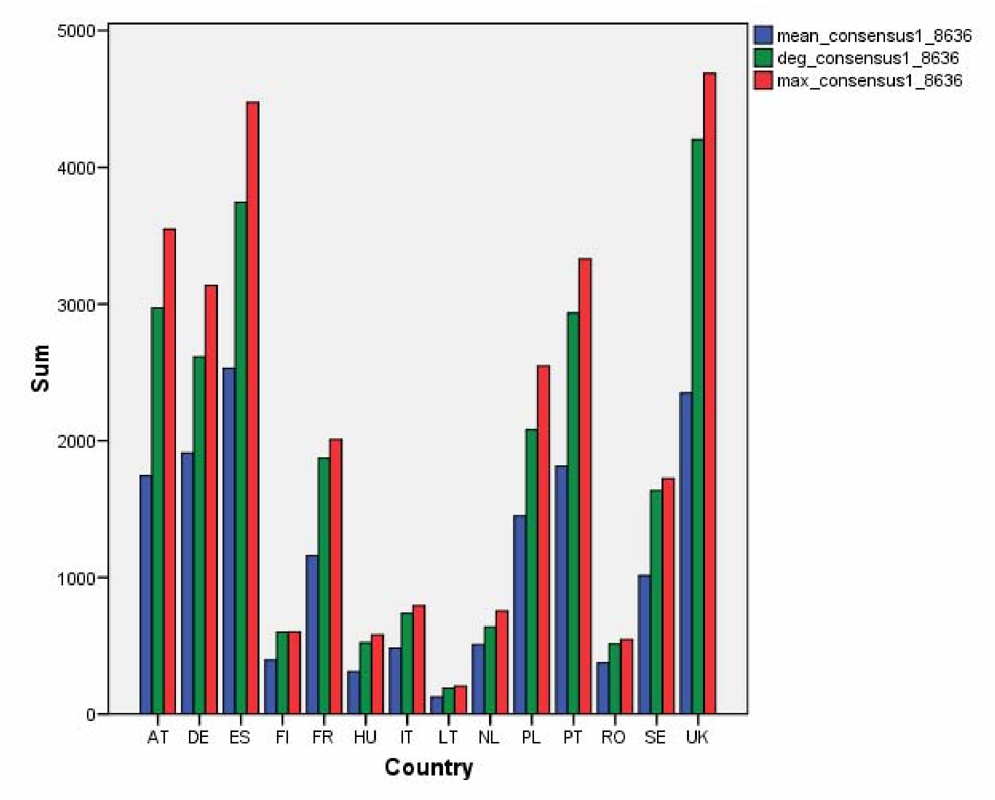

Figure 43: Summary of the 3 variants of indexes: arithmetic mean, degraded approach and worst case for the dataset consensus 1

Table 38: Valid, missing, mean und median values for all 3 variants of indexes for the dataset consensus 1

Statistics	mean_consensus1_8636	deg_consensus1_8636	max_consensus1_8636
Valid	8008	8266	8636
Missing	628	370	0
Mean	2.1	3.2	3.7
Median	2	4	4

Table 39: Frequency and percentage for the dataset consensus 1

	mean_consensus1_8636		deg_consensus1_8636		max_consensus1_8636	
Status	Frequency	%	Frequency	%	Frequency	%
1	1061	12.3	1461	16.9	1061	12.3
2	5441	63.0	28	0.3	399	4.6
3	1440	16.7	2432	28.2	2457	28.5
4	66	0.8	3941	45.6	1185	13.7
5			404	4.7	3534	40.9
Missing	628	7.3	370	4.3		
Total	8636	100	8636	100	8636	100

4. Variant, "Consensus 2" (9173) Sites

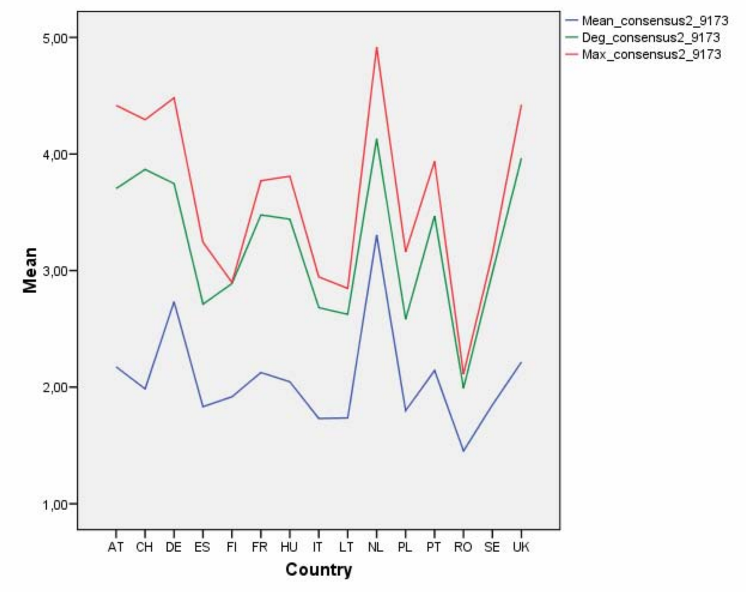

Figure 44: Distribution of the three different indices per country for the dataset consensus 2

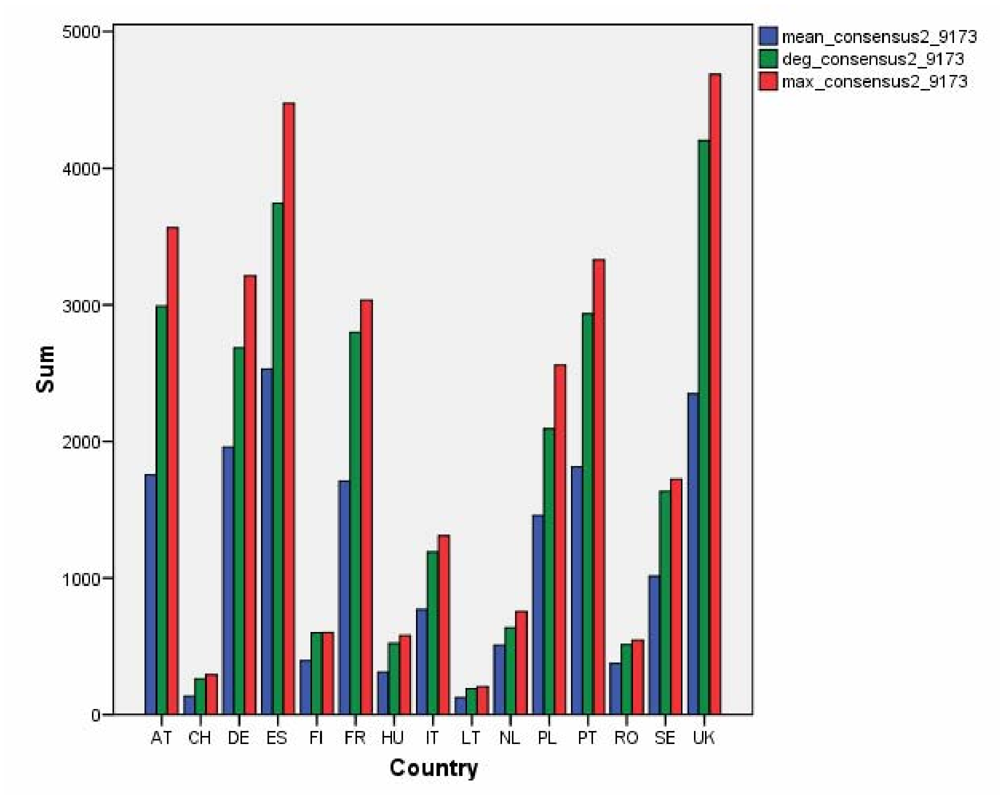

Figure 45: Summary of the 3 variants of indexes: arithmetic mean, degraded approach and worst case for the dataset consensus 2

Table 40: valid, missing, mean und median values for all 3 variants of indexes for the dataset consensus 2

Statistics	**mean_consensus2_9173**	**deg_consensus2_9173**	**max_consensus2_9173**
Valid	8513	8790	9173
Missing	660	383	0
Mean	2.1	3.2	3.7
Median	2	4	4

Table 41: Frequency and percentage for the for the dataset consensus 2

	mean_consensus2_9173		**max_consensus2_9173**		**deg_consensus2_9173**	
Status	Frequency	%	Frequency	%	Frequency	%
1	1100	12.0	1504	16.4	1100	12.0
2	5813	63.4	29	0.3	404	4.4
3	1534	16.7	2559	27.9	2585	28.2
4	66	0.7	4273	46.6	1327	14.5
5			425	4.6	3757	41.0
Missing	660	7.2	383	4.2		
Total	9173	100	9173	100	9173	100

5.Variant, "Minimum Data Loss" (9498 Sites)

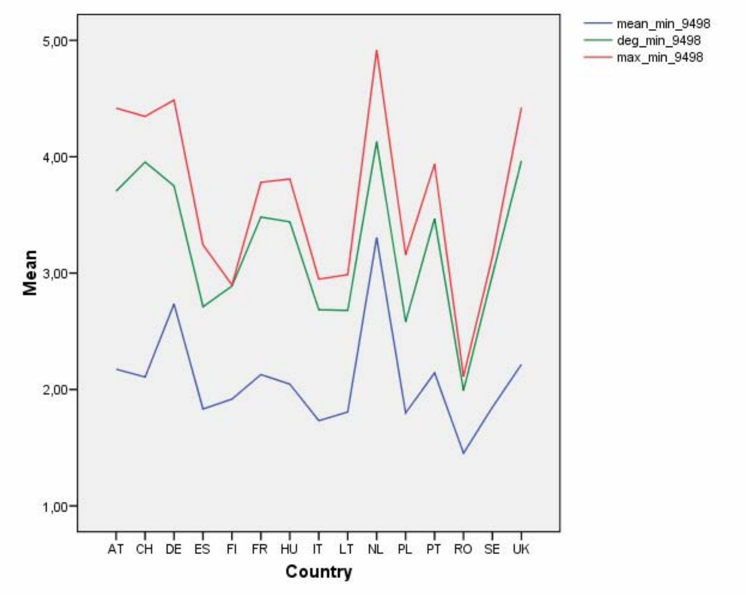

Figure 46: Distribution of the three different indices per country for the dataset minimum data loss

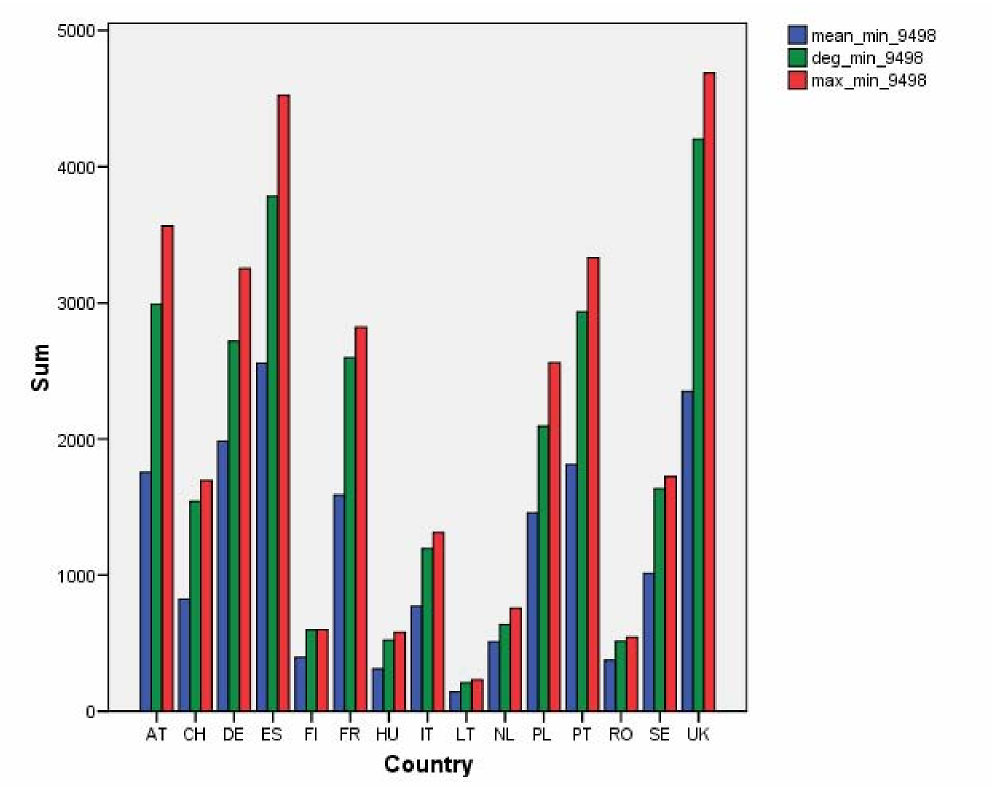

Figure 47: Summary of the 3 variants of indexes: arithmetic mean, degraded approach and worst case for the dataset minimum data loss

Table 42: Valid, missing, mean and median values all 3 Variants of indexes for the dataset minimum data loss

Statistics	mean_min_9498	deg_min__9498	max_ min__9498
Valid	8813	9098	9498
Missing	685	400	0
Mean	2.1	3.3	3.7
Median	2	4	4

Table 43: Frequency and percentage for the for the dataset minimum data loss

	mean_min__9498		deg_min__9498		max_min__9498	
Status	Frequency	%	Frequency	%	Frequency	%
1	1109	11.7	1518	16.0	1109	11.7
2	6043	63.6	29	0.3	416	4.4
3	1595	16.8	2579	27.2	2609	27.5
4	66	0.7	4490	47.3	1405	14.8
5			482	5.1	3959	41.7
Missing	685	7.2	400	4.2		
Total	9498	100	9498	100	9498	100

8. Discussion

8.1. Comparison of all Scenarios and Index Variants

In figure 48, the five index scenarios and the three index approaches are compared.

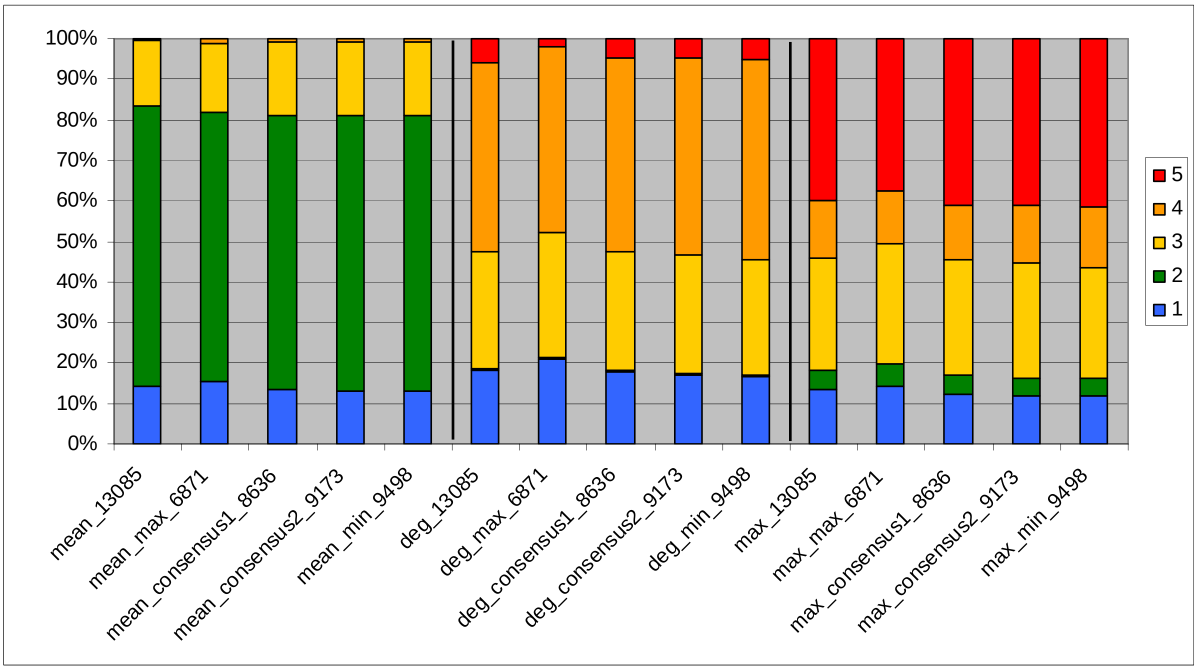

Figure 48: Comparison of all 5 scenarios and all 3 variants

Description for the abbreviations of figure 48:

Fame Approach:

mean 13085: FAME - Index Approach (arithmetic mean) of the dataset 13085 sites

mean_max_6871: FAME - Index Approach (arithmetic mean) of the index scenario maximum data loss (6871 sites)

mean_consensus1_8636: FAME - Index Approach (arithmetic mean) of the index scenario Consensus 1 (8636 sites)

mean_consensus2_9173: FAME - Index Approach (arithmetic mean) of the index scenario Consensus 2 (9173 sites)

mean_min_9498: FAME - Index Approach (arithmetic mean) of the index scenario Minimum data loss (9173 sites)

Degraded Approach:

Deg 13085: Degraded- Index Approach (average of all values worse then 2) of the dataset 13085 sites deg_max_6871: Degraded - Index Approach (average of all values worse then 2) of the index scenario maximum data loss (6871 sites)

deg_consensus1_8636: Degraded - Index Approach (average of all values worse then 2) of the index scenario Consensus1_8636

deg_consensus2_9173: Degraded - Index Approach (average of all values worse then 2) of the index scenario Consensus2_9173

deg_min_9498: Degraded - Index Approach (average of all values worse then 2) of the index scenario Minimum data loss (9498 sites)

Worst case approach:

max 13085: Worst Case - Index Approach (one out all out) of the dataset 13085 sites

max_max_6871: Worst Case - Index Approach (one out all out) of the index scenario maximum data loss (6871 sites)

max_consensus1_8636: Worst Case - Index Approach (one out all out) of the index scenario Consensus1_8636

max_consensus2_9173: Worst Case - Index Approach (one out all out) of the index scenario Consensus2_9173

max_min_9498: Worst Case - Index Approach (one out all out) of the index scenario Minimum data loss (9173 sites)

The EU - Water Framework Directive prescribes water quality status for the ecological river status assessment in the following steps: 1= high status, 2= good status, 3= moderate status, 4= poor status and 5=bad status.

It can be seen that for the different index scenarios are no major changes. Only the 3 index approaches give different values.

Variant 1: arithmetic mean ranges from high to moderate (1 to 3) whereby the status 1 and 2 are prevailing.

Variant 2: degraded approach represents the status from 1 till 5, where the status 3 and 4 are prevailing. The status 2 is not present.

Variant 3: worst case, shows status from 1 to 5, but the status 3 and 5 are prevailing.

8.2. Comparison of Pressure Data of the Data Set 13085 sites and the Data Set Consensus 1, (8636 sites)

The figures in this chapter show a comparison between all pressure groups and pressures between the dataset 13085 sites and consensus 1 (8636 sites). For the dataset 13085 sites all existing pressures were used. The amount and type of pressures used for the dataset Consensus 1 is explained in the table 33, Variants for Index Calculation. The abbreviations of the pressures are described in the table 32: Retaining variables for the Index development.

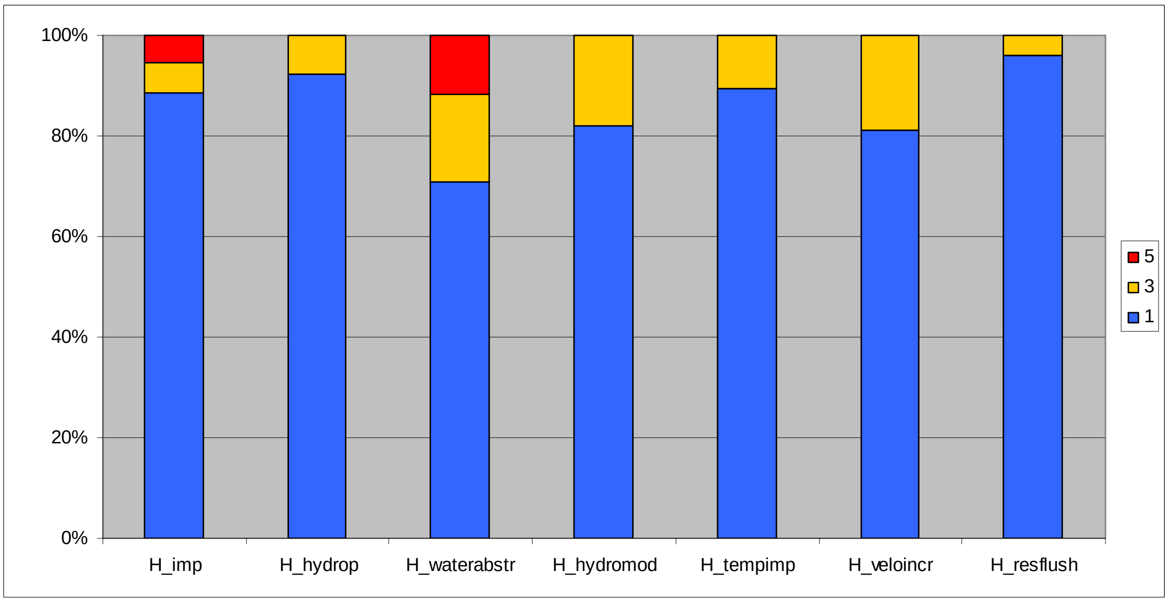

Figure 49: Summary of hydrological pressures, data set 13085 sites

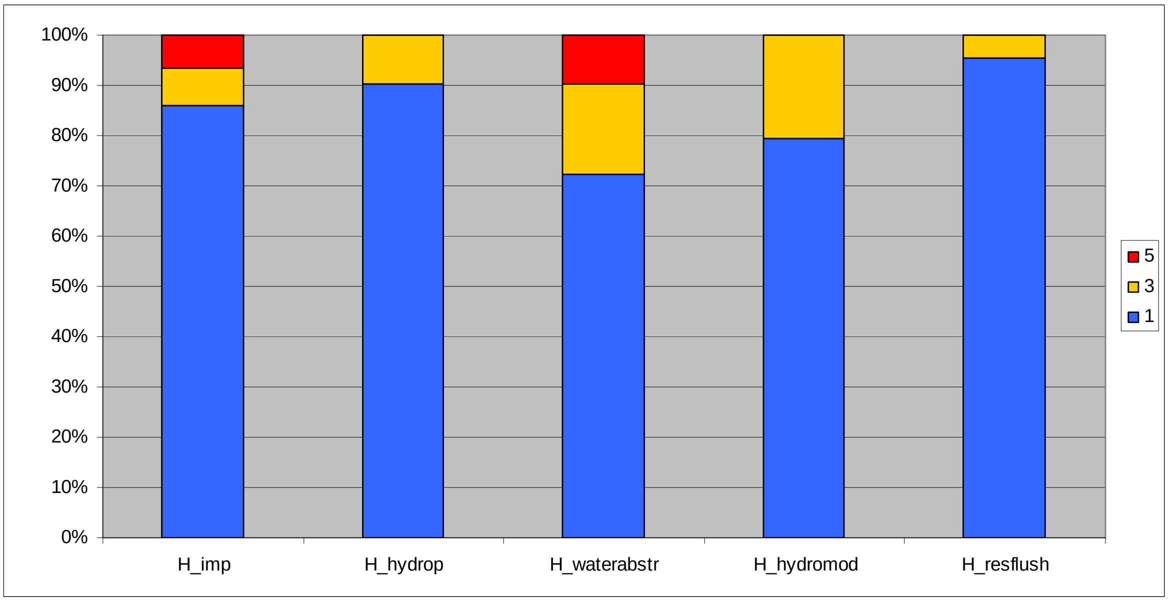

Figure 50: Summary of hydrological pressures, data set consensus 1 (8636 sites)

Between the hydrological pressure of the data dataset 13085 sites and Consensus 1 no differences can be found for the ratio of the different pressures.

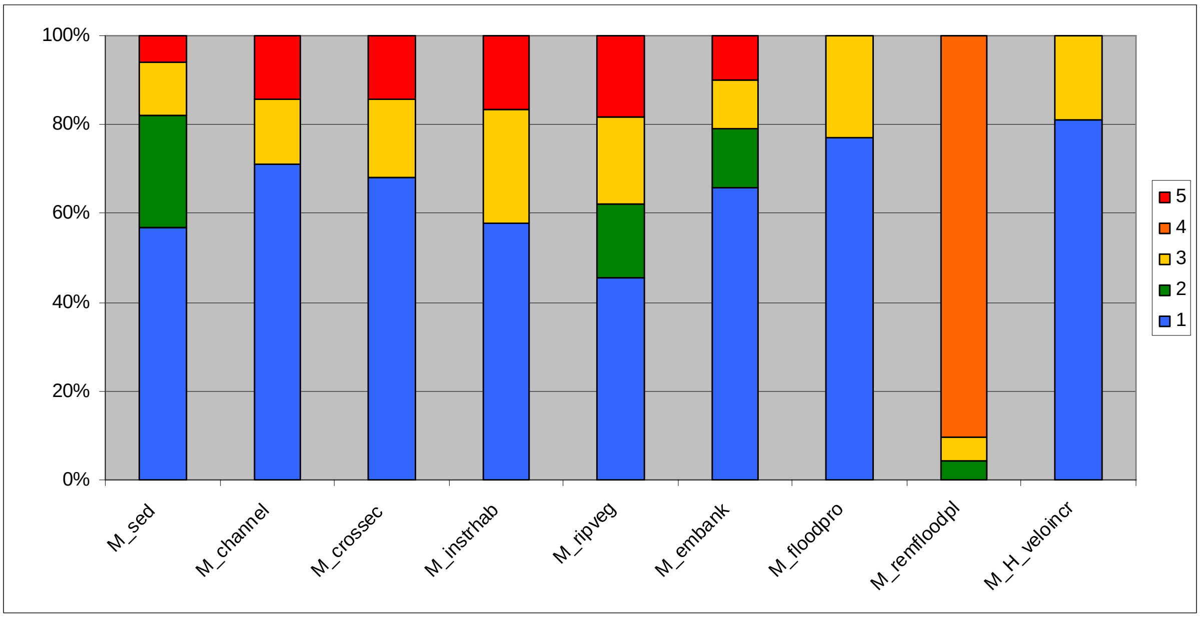

Figure 51: Summary of morphological pressures, data set 13085 sites

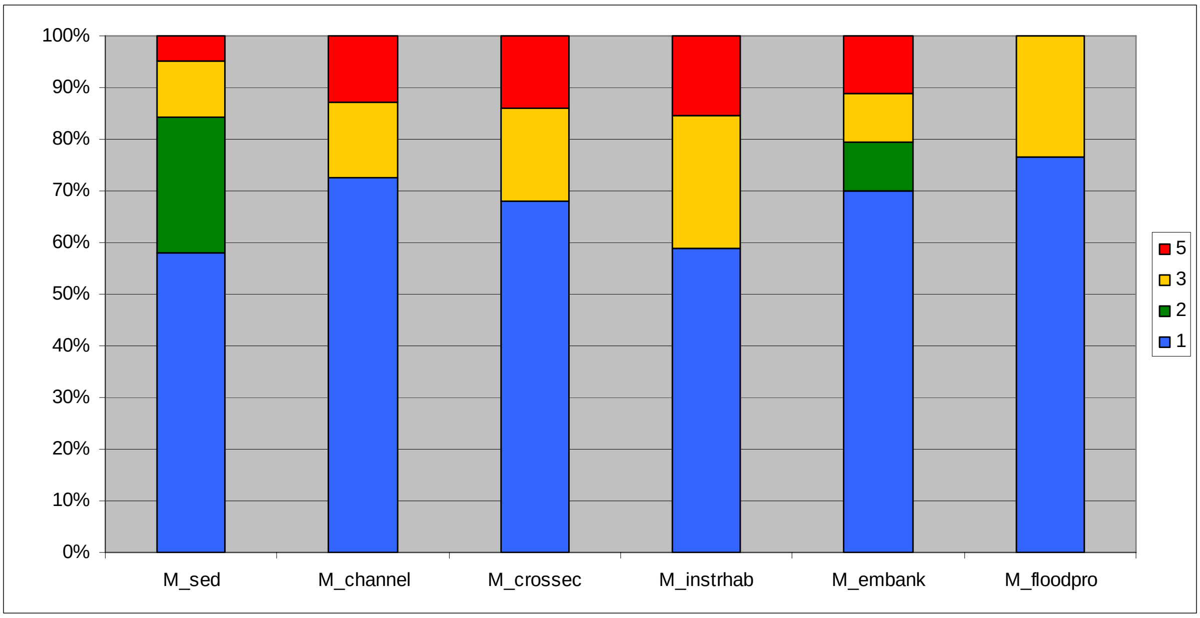

Figure 52: Summary of morphological pressures, data set consensus 1 (8636 sites)

Between the morphological pressures of the dataset 13085 sites and Consensus 1 no differences can be found for the ratio of the different pressures.

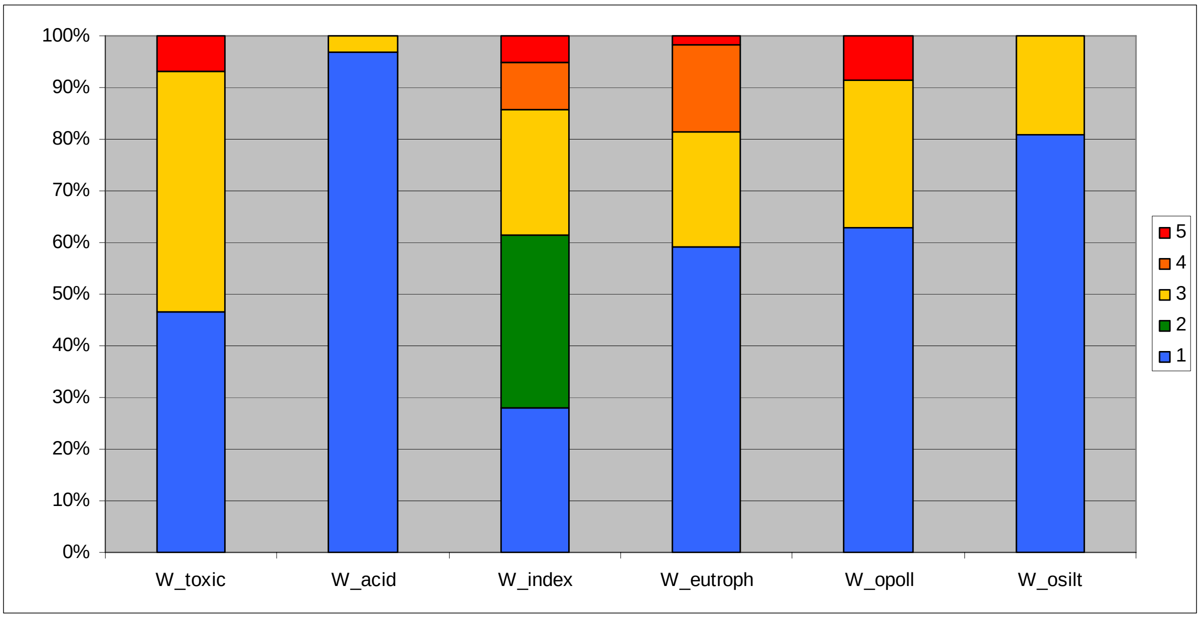

Figure 53: Summary of water quality pressures, data set 13085 sites

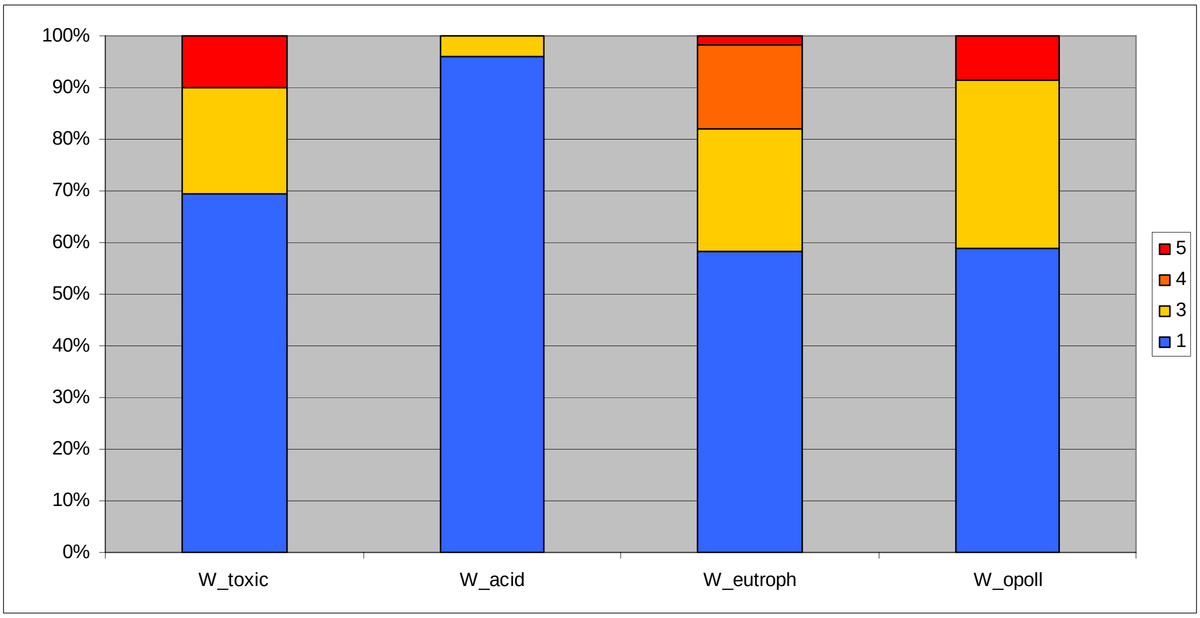

Figure 54: Summary of water quality pressures, data set consensus 1 (8636 sites)

Between the water quality pressures of the dataset 13085 sites and Consensus 1 no differences can be found for the ratio of the different pressures.

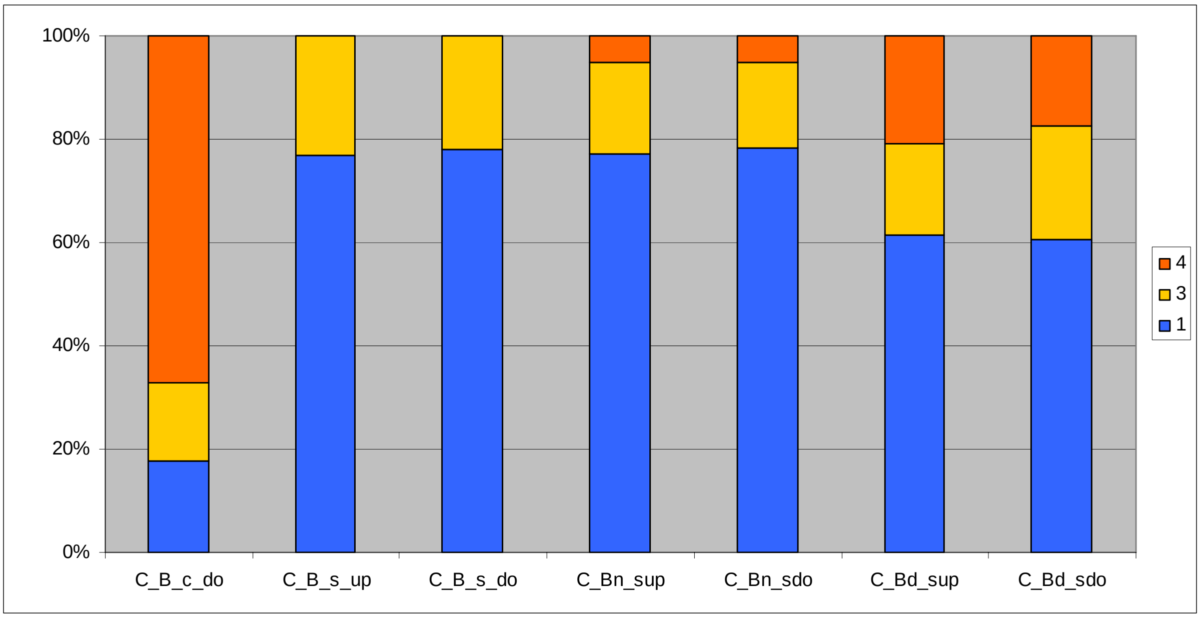

Figure 55: Summary of connectivity pressures, data set 13085 sites

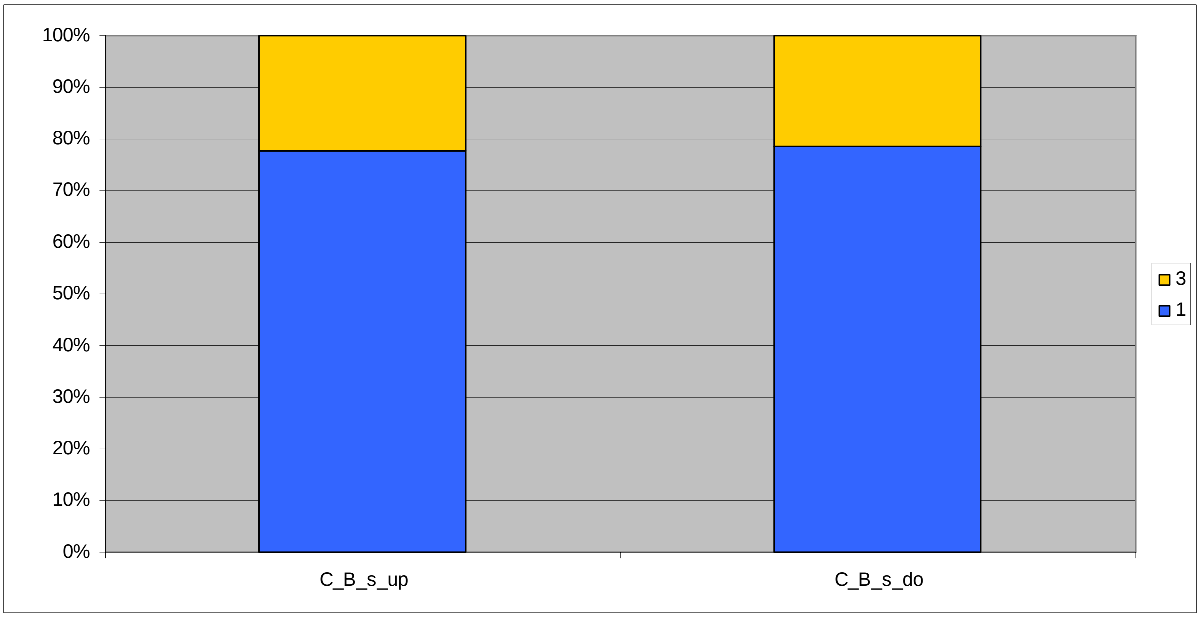

Figure 56: Summary of connectivity pressures, data set consensus 1 (8636 sites)

Between the connectivity pressures of the dataset 13085 sites and Consensus 1 no differences can be found for the ratio of the different pressures.

The tables 44 and 45 give an overview of the amount of data for every pressure variable and category.

Table 44: Overview over the amount of data of every pressure in the dataset 13085 sites

	H_imp	H_hydrop	H_waterabstr	H_hydromod	H_tempimp	H_veloincr	H_resflush		
1	10703	11837	8293	10388	9567	8477	12301		
3	752	984	2010	2299	1122	1978	497		
5	640		1384						
	M_sed	M_channel	M_crossec	M_instrhab	M_ripveg	M_embank	M_floodpro	M_remfloodpl	M_H_veloincr
1	6441	8596	8138	6912	4753	7708	9136		8477
2	2893				1756	1538		392	
3	1363	1786	2146	3053	2062	1281	2718	448	1978
4								7932	
5	665	1714	1691	1979	1902	1181			
	W_toxic	W_acid	W_index	W_eutroph	W_opoll	W_osilt			
1	7954	12373	3285	7137	7512	8236			
2			3934						
3	7954	393	2870	2701	3430	1942			
4			1067	2050					
5	1175		607	205	1024				
	C_B_c_do	C_B_s_up	C_B_s_do	C_Bn_sup	C_Bn_sdo	C_Bd_sup	C_Bd_sdo		
1	2287	9808	9952	8951	9029	3501	3374		
3	1919	2957	2802	2051	1891	1010	1233		
4	8615			612	605	1179	966		

Table 45: Overview over the amount of data of every pressure in the dataset consensus 1

	H_imp	H_hydrop	H_waterabstr	H_hydromod	H_resflush	
1	7420	7806	6249	6859	8252	
3	651	830	1539	1777	384	
5	562		848			
	M_sed	M_channel	M_crossec	M_instrhab	M_embank	M_floodpro
1	5002	6276	5868	5095	6035	6617
2	2271				813	
3	945	1254	1558	2203	826	2017
5	418	1105	1210	1338	960	
	W_toxic	W_acid	W_eutroph	W_opoll		
1	5980	8281	5043	5074		
3	1772	344	2050	2821		
4			1385			
5	862		158	730		
	C_B_s_up	C_B_s_do				
1	6721	6792				
3	1915	1843				

8.3. Comparison of the Index FAME and the EFI+ Project

8.4. The Index Development of the FAME Work Package 6/7

Melcher et al. 2003 described in the FAME paper, work package 6/7 spatial approach, the following 4 scenarios for an index development:

Scenario 1 – Five Main Variables:

For this approach, the five main FIDES variables for the human pressure group on rivers were chosen: Connectivity_segment, Hydrological_regime_site, Morphological_condition_site, Toxic acidification_site and Nutrients_organic_input_site.

Each fishing occasion was classified from 1 (high) to 5 (bad).

The impacts were unbalanced distributed. Toxic impact had only a class 1 and 2. Also nutrients impact had no classifications of 4 and 5.

Scenario 2 – Total Impact, Worst Case classification:

This “worst case classification” or also One out-all Out” is based on the fact that the worst impact class of one of the five variables gives the new value. This scenario is also used as the index scenario number 3 for this thesis.

Scenario 3 – Total Impact, Mean Values

This possibility to identify total impact classes is to compute the mean value between all five variables and afterwards the 5 classes grouped.

Scenario 4 – Total Impact, Sum of Scores:

All scores are summed up all and generate classes are created. In this special case two variations with five and with three groups were generated. This method is the same as using mean values.

The scenario is used also in this thesis, as scenario 1, FAME approach, principle: arithmetic mean.

8.5. Different Approaches of the Index Development in the Line with FAME :

Pont et. al. (2007) used for their approach 5252 sites and they also used only one fishing occasion per site.

This was done also for the EFI+ project were only the youngest date of the site was used so that the weighting of the data for each site is equal.

Ponts database contained 14 main variables to assess anthropogenic pressures at three scales: 1. basin, 2. river segment and 3. site. As only some of these variables were complete, only four of this 12 variables defined at the site scale were left over: These were: modification of morphology, hydrology, presence of toxic substances or acidification, and nutrient loading.

8.6. Index Development of the EFI+ Project for this Thesis

The dataset in this thesis is used in a different way. Four different scenarios were chosen by expert judgment: Max Data Loss, Consensus 1, Consensus 2 and Min Data Loss.

The approach for the Index development for this thesis is different. After expert judgment at the EFI+ consortium meeting in Lisbon, three different approaches of indexes were used.

The first approach was the same as it was used in the FAME: the arithmetic mean approach

As second approach the degraded approach was used. For this index a average of all values worse than 2 is calculated

The third approach is called the worst case approach. The principle for this approach is: “One out –all out” which means every time the highest count is used.

8.7. Metrics Development: Reference and Degraded Conditions

Schmutz et al. 2007 mentioned that pressure data are used for two purposes: to identify reference conditions and to calibrate indices. In the FAME project, reference conditions were developed using field data from only minimally or slightly impacted sites which means from sites which have the ecological high (1) and /or good (2) status.

In the figure 57 the blue line marks for all index scenarios the border of reference sites for the high status. As the WDF asks for the good status of the rivers, the high and good ecological status can be combined. This combination of status 1 and 2 is marked with the green line.

When for the 1. index approach, (arithmetic mean) all ecological high and good statuses are taken into account, almost 80 % of the dataset could be understand as a calibration site. The reason for this is that all sites transformed artificially with the calculation of the arithmetic mean to a better status.

For the index approach 2 and 3 only the high ecological status is available in the same amount. The reason for this is that only a few pressures variables have the category 2, which influences the low amount of sites of the good ecological status. Therefore, not enough calibration sites can be defined for the index approaches 2 and 3.

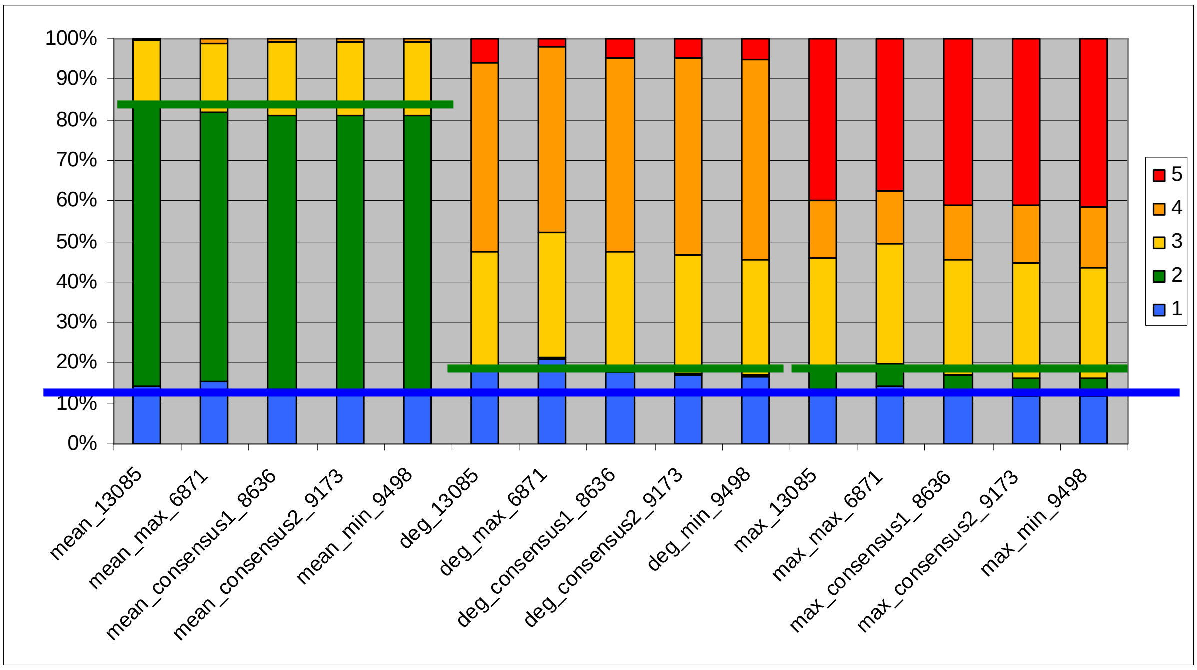

Figure 57: Comparison of all 5 scenarios and all 3 variants (borders for reference/calibration sites)

8.8. Proposal for an Index Development

When all 5 scenarios and 3 variants for the index development are analysed the following conclusions for a proposal of an index development can be drawn.

For all 5 scenarios of the first index approach (FAME, arithmetic mean) the figure 57 shows that the ecological high (1) and /or good (2) status of the sites are represented with 80 percent.

As it is totally unrealistic that all rivers in Europe have an ecological high and good status from over 80 percent this index approach can't be given any preference and will drop out.

The third index approach (worst case) shows that 40 percent of the sites are in bad ecological condition. Therefore, also this approach will drop out because this fact can be seen as very unrealistic.

The second index approach (degraded approach) shows the most balanced distribution of all ecological status, with the small exception that the good ecological status is not well represented. The reason for this is that as soon as a site is chemical or hydro morphological impacted the site is classified for this approach as moderate status. Another fact is that only a few pressures variables have the category 2. Nevertheless, this index approach can be seen as the favourable index approach for further index and metrics development.

For the further index and metrics development, the used dataset has to be without missing values. This approach narrows again the possibilities for a decision which scenario of the five available scenarios should be used in then future.

For this approach only the scenario "degraded approach minimal data" can be considered because of its complete dataset.

9. References

Breitholtz M., Hill C. & Bengtsson, B.-E. (2001) Toxic substances and reproductive disorders in Baltic fish and crustaceans. Ambio, Vol. 30, no. 4, pp. 210–216

Cowx I.G. & Gerdeaux D., (2004) The effects of fisheries management practises on freshwater ecosystems. Fish. Manage. Ecol. 11(3-4):145-151

Degerman E., Hammar J., Nyberg P. & Svärdson G. (2001) Human impact on the fish diversity in the four largest lakes of Sweden. Ambio 30(8):522-528

Degerman E., Beier U., Breine J., Melcher A., Quataert P.,Rogers C., Roset N. & Simoens I. (2007) Classification and assessment of degradation in European running waters.Fisheries Management and Ecology 14, 417–426

Doug A., Data Mining by Doug Alexander, (http://www.eco.utexas.edu/~norman/BUS.FOR/course.mat/Alex/), accessed in 07/2007

Davidson College; 2002, Department of Biology Davidson, NC 28036 (http://www.bio.davidson.edu/Courses/Bio111/Bio111LabMan/Lab%207.html), accessed in 07/2007

FAME CONSORTIUM (2004), Manual for the application of the European Fish Index - EFI. A Fish-based method to assess the ecological status of European rivers in support of the Water Framework Directive, Version 1.1, January 2005. (http://FAME.boku.ac.at)

Frawley W., and Piatetsky G., Shapiro and Matheus C. (Fall 1992). "Knowledge Discovery in Databases: An Overview". AI Magazine: pp. 213-228. ISSN 0738-4602

Galloway J.N. (2001), Acidification of the world: natural and anthropogenic. Water, Air, Soil Pollution. Vol. 130(1-4):17-24

Garcia-Berthou E. & Moreno-Amich R. (2000), Introduction of exotic fish into a Mediterranean lake over a 90-year period, Arch. Hydrobiol. 149(2):271-284.

Garner P., Bass, J.A.B., & Collett G.D., (1996), The effects of weed cutting upon the biota of a large regulated river. Aquat. Conserv.: Mar. Freshwat. Ecosyst. 6(1):21-29.

Hall L.W., Scott M.C., Killen W.D. Jr & Anderson R.D. (1996), The effects of land-use characteristics and acid sensitivity on the ecological status of Maryland coastal plain streams. Environ. Toxicol. Chem, 15(3):384-394

Hand D., Mannila H., Smyth P., (2001), Principles of Data Mining. MIT Press,Cambridge MA,. ISBN 0-262-08290-X

Hart B.T., Bailey P., Edwards R., Hortle K., James K., McMahon A., Meredith C. & Swadling, K., (1991), A review of the salt sensitivity of the Australian freshwater biota Hydrobiologia 210(1-2):105-144

Hernandez M.J. (2003), Database Design for Mere Mortals: A Hands-On Guide to Relational Database Design, 2nd edn. Boston: Addison-Wesley Professional, 611 pp.

Hlass L.J., Fisher W.L. & Turton D.J., (1998), Use of the index of biotic integrity to assess water quality in forested streams of the Quachita Mountains ecoregion, Arkansas. J. Freshwater Ecol. 13(2):181-192

Holcik J., (2003), Changes in the fish fauna and fisheries in the Slovak section of the Danube river: a review. Annales de limnologie 39(3):177-195

Improvement and Spatial Extension of the European Fish Index EFI+ 9, Newsletter March 2007

Jungwirth M., Haidvogl G., Moog, O., Muhar S., Schmutz S., (2003) Angewandte Fischökologie an Fleißgewässern Facultas Universitätsverlag Wien, UTB , ISBN:3-8252-2113-X

Karr J.R., (1981), Assessment of biotic integrity using fish communities. Fisheries 6 (6), 21–27

Lancaster H.O., An Introduction to Medical Statistics. John Wiley and Sons. 1974. ISBN 0 471 51250-8

Lessard J.L. & Hayes D.B., (2003), Effects of elevated water temperatures on fish and macroinvertebrates below small dams. River Res. Appl. 19(7):721-732

Melcher A., Schmutz S., & Haidvogl G., FAME WP6/7 SPATIAL APPROACH ALPS/ECOREGION 4 REFERENCE AND DEGRADED CONDITIONS

Melcher A., Schmutz S., Haidvogl G. & Moder K., (2007), Spatially based methods to assess the ecological status of European fish assemblage types. Fisheries Management and Ecology 14, 453–463

Pont D., Hugueny B., Beier U., Goffaux D., Melcher A., Noble R., Rogers C., Roset N. & Schmutz S. (2006), Assessing the biotic integrity of rivers at the continental scale: a European approach using fish assemblages. Journal of Applied Ecology, 43, 70-80

Pont D., Hugueny B. & Rogers C., (2007), Development of a fish-based index for the assess

Schmutz S., Cowx I.G., Haidvogl G., & Pont D. (2007), Fish-based methods for assessing European running waters: a synthesis

Schmutz S., Melcher A., Muhar S., Zitek A., Poppe M., Trautwein C., & Jungwirth M., (2007), MIRR- A Model-based instrument for River Restoration. Entwicklung eines strategischen Instruments zur integrativen Bewertung ökologischer Restaurationsmaßnahmen an Fließgewässern

Stanford J.A. and Ward J.V. (2001). "Revisiting the serial discontinuity concept." Regulated Rivers-Research & Management 17(4-5): 303-310

The European Fish Index, Fisheries Management and Ecology 14, 427–439.

Wang L., Lyons J., Kanehl P., Bannerman R. & Emmons E., (2000), Watershed urbanization and changes in fish communities in southeastern Wisconsin streams. J. Am. Water Resour. Assoc. 36(5):1173-1175

WFD (2000), DIRECTIVE 2000/60/EC OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 23 October 2000 establishing a framework for Community action in the field of water policy

Zalewski M., Schiemer F. & Thorpe J. (2001), Fish and land-inland ecotones – overview and synthesis. Int. J. Ecohydrology Hydrobiol. 1(1-2):261-266

Zitek A., Santocildes G., Schmutz W., Schmutz S. (2006), “POTENTIAL CRITERIA FOR MODELLING FISH/PRESSURE RELATIONSHIPS IN RUNNING WATERS A LITERATURE REVIEW”, Institute of Hydrobiology and Aquatic Ecosystem Management, Department of Water, Atmosphere and Environment, BOKU - University of Natural Resources and Applied Life Sciences, Vienna prepared for The Austrian Federal

Ministry of Agriculture, Forestry, Environment and Water Management & The Government of Lower Austria. http://mirr.boku.ac.at/mirr_resultate.htm

Internet References:

The FAME project, http://fame.boku.ac.at/

The MIRR project, http://mirr.boku.ac.at/mirr_resultate.htm

The EFI+ project, http://efi-plus.boku.ac.at/

A Beginner's Guide For Descriptive Statistics, Department of Biology, Davidson College, Davidson, NC 28036, (http://www.bio.davidson.edu/Courses/Bio111/Bio111LabMan/Lab%207.html), accessed in 07/2007

Bar Graph:Wikipedia, (http://en.wikipedia.org/wiki/Bar_graph), accessed in 07/2007

Cluster Analysis: Wikipedia, (http://en.wikipedia.org/wiki/Cluster_analysis), accessed in 07/2007

Discriminant analyses, (http://www.eso.org/sci/data processing/software/esomidas//doc/user/98NOV/volb/node214.html), accessed in 02/2008

ESO, 1997, European Organisation for Astronomical Research in the Southern Hemisphere,(http://www.eso.org/sci/dataprocessing/software/esomidas//doc/user/98NOV/volb/node214.html), accessed in 02/2008

Factor Analyses: (http://en.wikipedia.org/wiki/Factor_analysis, Richard B. Darlington http://comp9.psych.cornell.edu/Darlington/factor.htm), accessed in 07/2007

Histogram: Wikipedia, (http://en.wikipedia.org/wiki/Histogram), accessed in 02/2008

Spearman's Rank Correlation: RevisionNotes.Co.Uk, (http://www.revision-notes.co.uk/revision/181.html, accessed in 01/2008 “Water Analysis – Fishing with Electricity (EN 14011; CEN, 2003)), accessed in 07/2007

10. Annexes

The annexes consists of detailed tables of all pressures with the there missing values and values for all categories for the dataset N=13 085 sites.

Hydrological Pressures

Table 46: Distribution of impoundment

Impoundment						
Country		Missing	No	Weak	Strong	Total
AT	Count	1.0	759.0	55.0	107.0	922
	%	0.1	82.3	6.0	11.6	100
CH	Count	0.0	717.0	0.0	0.0	717
	%	0.0	100.0	0.0	0.0	100
DE	Count	0.0	761.0	23.0	1.0	785
	%	0.0	96.9	2.9	0.1	100
ES	Count	0.0	3125.0	12.0	10.0	3147
	%	0.0	99.3	0.4	0.3	100
FI	Count	**226.0**	304.0	0.0	0.0	530
	%	**42.6**	57.4	0.0	0.0	100
FR	Count	38.0	906.0	107.0	94.0	1145
	%	3.3	79.1	9.3	8.2	100
HU	Count	0.0	188.0	5.0	0.0	193
	%	0.0	97.4	2.6	0.0	100
IT	Count	0.0	610.0	16.0	26.0	652
	%	0.0	93.6	2.5	4.0	100
LT	Count	0.0	109.0	2.0	4.0	115
	%	0.0	94.8	1.7	3.5	100
NL	Count	0.0	23.0	43.0	116.0	182
	%	0.0	12.6	23.6	63.7	100
PL	Count	0.0	805.0	83.0	31.0	919
	%	0.0	87.6	9.0	3.4	100
PT	Count	0.0	698.0	138.0	87.0	923
	%	0.0	75.6	15.0	9.4	100
RO	Count	0.0	228.0	7.0	28.0	263
	%	0.0	86.7	2.7	10.6	100
SE	Count	0.0	472.0	86.0	47.0	605
	%	0.0	78.0	14.2	7.8	100
UK	Count	**723.0**	999.0	174.0	91.0	1987
	%	**36.4**	50.3	8.8	4.6	100
Total	Count	988.0	10,704.0	751.0	642.0	13,085
	%	7.6	81.8	5.7	4.9	100

Table 47: Distribution of hydropeaking

Hydropeaking						
Country		Missing	No	Partial	Yes	Total
AT	Count	4.0	855.0	0.0	63.0	922.0
	%	0.4	92.7	0.0	6.8	100.0
CH	Count	0.0	704.0	0.0	13.0	717.0
	%	0.0	98.2	0.0	1.8	100.0
DE	Count	0.0	752.0	0.0	33.0	785.0
	%	0.0	95.8	0.0	4.2	100.0
ES	Count	0.0	2993.0	83.0	71.0	3147.0
	%	0.0	95.1	2.6	2.3	100.0
FI	Count	**226.0**	297.0	0.0	7.0	530.0
	%	**42.6**	56.0	0.0	1.3	100.0
FR	Count	30.0	1008.0	0.0	107.0	1145.0
	%	2.6	88.0	0.0	9.3	100.0
HU	Count	0.0	192.0	0.0	1.0	193.0
	%	0.0	99.5	0.0	0.5	100.0
IT	Count	0.0	608.0	11.0	33.0	652.0
	%	0.0	93.3	1.7	5.1	100.0
LT	Count	1.0	109.0	1.0	4.0	115.0
	%	0.9	94.8	0.9	3.5	100.0
NL	Count	0.0	178.0	0.0	4.0	182.0
	%	0.0	97.8	0.0	2.2	100.0
PL	Count	3.0	869.0	18.0	29.0	919.0
	%	0.3	94.6	2.0	3.2	100.0
PT	Count	0.0	542.0	88.0	293.0	923.0
	%	0.0	58.7	9.5	31.7	100.0
RO	Count	0.0	245.0	0.0	18.0	263.0
	%	0.0	93.2	0.0	6.8	100.0
SE	Count	0.0	497.0	63.0	45.0	605.0
	%	0.0	82.1	10.4	7.4	100.0
UK	Count	0.0	1987.0	0.0	0.0	1987.0
	%	0.0	100.0	0.0	0.0	100.0
Total	Count	264.0	11,836.0	264.0	721.0	13085.0
	%	2.0	90.5	2.0	5.5	100.0

Table 48: Distribution of water_abstraction

Country		Water_abstraction Missing	No	Weak	Strong	Total
AT	Count	7.0	733.0	80.0	102.0	922.0
	%	0.8	79.5	8.7	11.1	100.0
CH	Count	0.0	532.0	51.0	134.0	717.0
	%	0.0	74.2	7.1	18.7	100.0
DE	Count	0.0	349.0	428.0	8.0	785.0
	%	0.0	44.5	54.5	1.0	100.0
ES	Count	**1013.0**	1622.0	230.0	282.0	3147.0
	%	**32.2**	51.5	7.3	9.0	100.0
FI	Count	**226.0**	302.0	1.0	1.0	530.0
	%	**42.6**	57.0	0.2	0.2	100.0
FR	Count	63.0	621.0	351.0	110.0	1145.0
	%	5.5	54.2	30.7	9.6	100.0
HU	Count	0.0	171.0	21.0	1.0	193.0
	%	0.0	88.6	10.9	0.5	100.0
IT	Count	88.0	291.0	154.0	119.0	652.0
	%	13.5	44.6	23.6	18.3	100.0
LT	Count	0.0	63.0	39.0	13.0	115.0
	%	0.0	54.8	33.9	11.3	100.0
NL	Count	0.0	49.0	8.0	125.0	182.0
	%	0.0	26.9	4.4	68.7	100.0
PL	Count	0.0	828.0	54.0	37.0	919.0
	%	0.0	90.1	5.9	4.0	100.0
PT	Count	0.0	480.0	317.0	126.0	923.0
	%	0.0	52.0	34.3	13.7	100.0
RO	Count	0.0	241.0	9.0	13.0	263.0
	%	0.0	91.6	3.4	4.9	100.0
SE	Count	0.0	561.0	27.0	17.0	605.0
	%	0.0	92.7	4.5	2.8	100.0
UK	Count	0.0	1453.0	239.0	295.0	1987.0
	%	0.0	73.1	12.0	14.8	100.0
Total	Count	1397.0	8296.0	2009.0	1383.0	13,085.0
	%	10.7	63.4	15.4	10.6	100.0

Table 49: Table: Distribution of water_use

Water_use													
Country		Missing	Cooling	Drink. water	Fish ponds	Hydropower	Indust.	Irrig.	No	Other	Snow P.	Tourism Industry	Total
AT	Count	0.0	0.0	0.0	0.0	381.0	0.0	0.0	541.0	0.0	0.0	0.0	922.0
	%	0.0	0.0	0.0	0.0	41.3	0.0	0.0	58.7	0.0	0.0	0.0	100.0
CH	Count	0.0	0.0	0.0	0.0	180.0	0.0	0.0	536.0	1.0	0.0	0.0	717.0
	%	0.0	0.0	0.0	0.0	25.1	0.0	0.0	74.8	0.1	0.0	0.0	100.0
DE	Count	0.0	0.0	0.0	4.0	1.0	276.0	145.0	349.0	10.0	0.0	0.0	785.0
	%	0.0	0.0	0.0	0.5	0.1	35.2	18.5	44.5	1.3	0.0	0.0	100.0
ES	Count	**977.0**	0.0	190.0	0.0	179.0	1.0	201.0	1573.0	26.0	0.0	0.0	3147.0
	%	**31.0**	0.0	6.0	0.0	5.7	0.0	6.4	50.0	0.8	0.0	0.0	100.0
FI	Count	**226.0**	0.0	0.0	0.0	15.0	0.0	0.0	287.0	2.0	0.0	0.0	530.0
	%	**42.6**	0.0	0.0	0.0	2.8	0.0	0.0	54.2	0.4	0.0	0.0	100.0
FR	Count	79.0	7.0	102.0	17.0	62.0	0.0	221.0	618.0	0.0	2.0	37.0	1145.0
	%	6.9	0.6	8.9	1.5	5.4	0.0	19.3	54.0	0.0	0.2	3.2	100.0
HU	Count	0.0	0.0	4.0	8.0	2.0	3.0	18.0	142.0	16.0	0.0	0.0	193.0
	%	0.0	0.0	2.1	4.1	1.0	1.6	9.3	73.6	8.3	0.0	0.0	100.0
IT	Count	90.0	2.0	28.0	12.0	97.0	2.0	129.0	290.0	2.0	0.0	0.0	652.0
	%	13.8	0.3	4.3	1.8	14.9	0.3	19.8	44.5	0.3	0.0	0.0	100.0
LT	Count	0.0	0.0	3.0	6.0	14.0	2.0	20.0	63.0	7.0	0.0	0.0	115.0
	%	0.0	0.0	2.6	5.2	12.2	1.7	17.4	54.8	6.1	0.0	0.0	100.0
NL	Count	0.0	35.0	0.0	0.0	0.0	0.0	147.0	0.0	0.0	0.0	0.0	182.0
	%	0.0	19.2	0.0	0.0	0.0	0.0	80.8	0.0	0.0	0.0	0.0	100.0
PL	Count	0.0	0.0	15.0	68.0	19.0	9.0	24.0	784.0	0.0	0.0	0.0	919.0
	%	0.0	0.0	1.6	7.4	2.1	1.0	2.6	85.3	0.0	0.0	0.0	100.0
PT	Count	0.0	0.0	0.0	0.0	358.0	0.0	98.0	467.0	0.0	0.0	0.0	923.0
	%	0.0	0.0	0.0	0.0	38.8	0.0	10.6	50.6	0.0	0.0	0.0	100.0
RO	Count	0.0	2.0	1.0	0.0	21.0	4.0	4.0	231.0	0.0	0.0	0.0	263.0
	%	0.0	0.8	0.4	0.0	8.0	1.5	1.5	87.8	0.0	0.0	0.0	100.0
SE	Count	0.0	0.0	0.0	0.0	20.0	2.0	22.0	561.0	0.0	0.0	0.0	605.0
	%	0.0	0.0	0.0	0.0	3.3	0.3	3.6	92.7	0.0	0.0	0.0	100.0
UK	Count	**1987.0**	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1987.0
	%	**100.0**	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	100.0
Total	Count	3359.0	46.0	343.0	115.0	1349.0	299.0	1029.0	6442.0	64.0	2.0	37.0	13, 085.0
	%	25.7	0.4	2.6	0.9	10.3	2.3	7.9	49.2	0.5	0.0	0.3	100.0

Table 50: Distribution of hydro_mod

Hydro_mod					
Country		missing	No	Yes	Total
AT	Count	0.0	894.0	28.0	922.0
	%	0.0	97.0	3.0	100.0
CH	Count	0.0	714.0	3.0	717.0
	%	0.0	99.6	0.4	100.0
DE	Count	0.0	633.0	152.0	785.0
	%	0.0	80.6	19.4	100.0
ES	Count	0.0	2932.0	215.0	3147.0
	%	0.0	93.2	6.8	100.0
FI	Count	**226.0**	287.0	17.0	530.0
	%	**42.6**	54.2	3.2	100.0
FR	Count	52.0	664.0	429.0	1145.0
	%	4.5	58.0	37.5	100.0
HU	Count	0.0	150.0	43.0	193.0
	%	0.0	77.7	22.3	100.0
IT	Count	89.0	511.0	52.0	652.0
	%	13.7	78.4	8.0	100.0
LT	Count	20.0	54.0	41.0	115.0
	%	17.4	47.0	35.7	100.0
NL	Count	0.0	53.0	129.0	182.0
	%	0.0	29.1	70.9	100.0
PL	Count	4.0	814.0	101.0	919.0
	%	0.4	88.6	11.0	100.0
PT	Count	0.0	403.0	520.0	923.0
	%	0.0	43.7	56.3	100.0
RO	Count	0.0	242.0	21.0	263.0
	%	0.0	92.0	8.0	100.0
SE	Count	0.0	468.0	137.0	605.0
	%	0.0	77.4	22.6	100.0
UK	Count	7.0	1564.0	416.0	1987.0
	%	0.4	78.7	20.9	100.0
Total	Count	398.0	10,383.0	2304.0	13,085.0
	%	3.0	79.4	17.6	100.0

Table 51: Distribution of temperature_impact

Temperature_impact						
Country		Missing	No	Permanent decrease	Permanent increase	Summer decrease
AT	Count	24.0	898.0	0.0	0.0	0.0
	%	2.6	97.4	0.0	0.0	0.0
CH	Count	0.0	583.0	0.0	0.0	0.0
	%	0.0	81.3	0.0	0.0	0.0
DE	Count	0.0	508.0	0.0	274.0	0.0
	%	0.0	64.7	0.0	34.9	0.0
ES	Count	0.0	2944.0	0.0	0.0	73.0
	%	0.0	93.5	0.0	0.0	2.3
FI	Count	**226.0**	304.0	0.0	0.0	0.0
	%	**42.6**	57.4	0.0	0.0	0.0
FR	Count	66.0	814.0	4.0	31.0	26.0
	%	5.8	71.1	0.3	2.7	2.3
HU	Count	0.0	192.0	0.0	1.0	0.0
	%	0.0	99.5	0.0	0.5	0.0
IT	Count	69.0	452.0	0.0	4.0	9.0
	%	10.6	69.3	0.0	0.6	1.4
LT	Count	18.0	88.0	0.0	0.0	0.0
	%	15.7	76.5	0.0	0.0	0.0
NL	Count	0.0	147.0	0.0	35.0	0.0
	%	0.0	80.8	0.0	19.2	0.0
PL	Count	6.0	888.0	0.0	12.0	6.0
	%	0.7	96.6	0.0	1.3	0.7
PT	Count	0.0	911.0	11.0	1.0	0.0
	%	0.0	98.7	1.2	0.1	0.0
RO	Count	0.0	254.0	0.0	0.0	9.0
	%	0.0	96.6	0.0	0.0	3.4
SE	Count	0.0	583.0	0.0	0.0	0.0
	%	0.0	96.4	0.0	0.0	0.0
UK	Count	**1987.0**	0.0	0.0	0.0	0.0
	%	**100.0**	0.0	0.0	0.0	0.0
Total	Count	2396.0	9566.0	15.0	358.0	123.0
	%	18.3	73.1	0.1	2.7	0.9

Table 52: Distribution of velocity_increase

Velocity_increase					
Country		Missing	No	Yes	Total
AT	Count	2.0	715.0	205.0	922.0
	%	0.2	77.5	22.2	100.0
CH	Count	0.0	319.0	398.0	717.0
	%	0.0	44.5	55.5	100.0
DE	Count	0.0	234.0	551.0	785.0
	%	0.0	29.8	70.2	100.0
ES	Count	263.0	2548.0	336.0	3147.0
	%	8.4	81.0	10.7	100.0
FI	Count	**226.0**	302.0	2.0	530.0
	%	**42.6**	57.0	0.4	100.0
FR	Count	43.0	1007.0	95.0	1145.0
	%	3.8	87.9	8.3	100.0
HU	Count	0.0	173.0	20.0	193.0
	%	0.0	89.6	10.4	100.0
IT	Count	83.0	483.0	86.0	652.0
	%	12.7	74.1	13.2	100.0
LT	Count	24.0	91.0	0.0	115.0
	%	20.9	79.1	0.0	100.0
NL	Count	0.0	42.0	140.0	182.0
	%	0.0	23.1	76.9	100.0
PL	Count	3.0	875.0	41.0	919.0
	%	0.3	95.2	4.5	100.0
PT	Count	0.0	851.0	72.0	923.0
	%	0.0	92.2	7.8	100.0
RO	Count	0.0	232.0	31.0	263.0
	%	0.0	88.2	11.8	100.0
SE	Count	0.0	605.0	0.0	605.0
	%	0.0	100.0	0.0	100.0
UK	Count	**1987.0**	0.0	0.0	1987.0
	%	**100.0**	0.0	0.0	100.0
Total	Count	2631.0	8477.0	1977.0	13,085.0
	%	20.1	64.8	15.1	100.0

Table 53: Distribution of reservoir_flushing

Reservoir_flushing					
Country		Missing	No	Yes	Total
AT	Count	33.0	889.0	0.0	922.0
	%	3.6	96.4	0.0	100.0
CH	Count	0.0	702.0	15.0	717.0
	%	0.0	97.9	2.1	100.0
DE	Count	0.0	620.0	165.0	785.0
	%	0.0	79.0	21.0	100.0
ES	Count	0.0	3080.0	67.0	3147.0
	%	0.0	97.9	2.1	100.0
FI	Count	**226.0**	304.0	0.0	530.0
	%	**42.6**	57.4	0.0	100.0
FR	Count	27.0	1066.0	52.0	1145.0
	%	2.4	93.1	4.5	100.0
HU	Count	0.0	193.0	0.0	193.0
	%	0.0	100.0	0.0	100.0
IT	Count	0.0	619.0	33.0	652.0
	%	0.0	94.9	5.1	100.0
LT	Count	0.0	112.0	3.0	115.0
	%	0.0	97.4	2.6	100.0
NL	Count	0.0	67.0	115.0	182.0
	%	0.0	36.8	63.2	100.0
PL	Count	1.0	887.0	31.0	919.0
	%	0.1	96.5	3.4	100.0
PT	Count	0.0	923.0	0.0	923.0
	%	0.0	100.0	0.0	100.0
RO	Count	0.0	248.0	15.0	263.0
	%	0.0	94.3	5.7	100.0
SE	Count	0.0	605.0	0.0	605.0
	%	0.0	100.0	0.0	100.0
UK	Count	0.0	1987.0	0.0	1987.0
	%	0.0	100.0	0.0	100.0
Total	Count	287.0	12302.0	496.0	13,085.0
	%	2.2	94.0	3.8	100.0

Water Quality Pressures

Table 54: Distribution of water_quality_index

Water_quality_index							
Country		1	2	3	4	5	Total
AT	Count	316.0	596.0	10.0	0.0	0.0	922.0
	%	34.3	64.6	1.1	0.0	0.0	100.0
CH	Count	241.0	169.0	130.0	98.0	75.0	713.0
	%	33.8	23.7	18.2	13.7	10.5	100.0
DE	Count	1.0	64.0	720.0	0.0	0.0	785.0
	%	0.1	8.2	91.7	0.0	0.0	100.0
ES	Count	1295.0	930.0	311.0	125.0	90.0	2751.0
	%	47.1	33.8	11.3	4.5	3.3	100.0
FI	Count	65.0	135.0	73.0	36.0	1.0	310.0
	%	21.0	43.5	23.5	11.6	0.3	100.0
FR	Count	114.0	334.0	155.0	55.0	26.0	684.0
	%	16.7	48.8	22.7	8.0	3.8	100.0
HU	Count	3.0	47.0	85.0	56.0	2.0	193.0
	%	1.6	24.4	44.0	29.0	1.0	100.0
IT	Count	169.0	194.0	102.0	25.0	2.0	492.0
	%	34.3	39.4	20.7	5.1	0.4	100.0
LT	Count	27.0	42.0	39.0	5.0	2.0	115.0
	%	23.5	36.5	33.9	4.3	1.7	100.0
NL	Count	0.0	102.0	79.0	1.0	0.0	182.0
	%	0.0	56.0	43.4	0.5	0.0	100.0
PL	Count	54.0	286.0	385.0	138.0	49.0	912.0
	%	5.9	31.4	42.2	15.1	5.4	100.0
PT	Count	0.0	8.0	265.0	373.0	277.0	923.0
	%	0.0	0.9	28.7	40.4	30.0	100.0
RO	Count	177.0	58.0	18.0	6.0	4.0	263.0
	%	67.3	22.1	6.8	2.3	1.5	100.0
SE	Count	37.0	296.0	154.0	52.0	27.0	566.0
	%	6.5	52.3	27.2	9.2	4.8	100.0
UK	Count	775.0	721.0	342.0	90.0	44.0	1972.0
	%	39.3	36.6	17.3	4.6	2.2	100.0
Total	Count	3274.0	3982.0	2868.0	1060.0	599.0	11,783.0
	%	27.8	33.8	24.3	9.0	5.1	100.0

Table 55: Distribution of water_quality_name

		Water_quality_name																								
		Missing	Bad	CCAS	Chem. GQA	El chepa	exp.	exp. FWA	Extreme	Feasibility	former classificatio	FWA	Good	High	IBE	Inland Surface Water	Intermediate	LAWA	Low	Moderate	No	Pantle_ Buck saprobic	Poor	Saprobienindex	VIX	Total
AT	Count	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	922	0	922
	%	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	100	0	100
CH	Count	4	0	0	0	0	0	0	173	0	0	0	0	0	0	0	130	0	169	0	241	0	0	0	0	717
	%	1	0	0	0	0	0	0	24	0	0	0	0	0	0	0	18	0	24	0	34	0	0	0	0	100
DE	Count	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	785	0	0	0	0	0	0	0	785
	%	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	100	0	0	0	0	0	0	0	100
ES	Count	**396**	65	0	0	987	0	0	0	0	0	0	616	786	0	0	0	0	0	176	0	0	121	0	0	3147
	%	**13**	2	0	0	31	0	0	0	0	0	0	20	25	0	0	0	0	0	6	0	0	4	0	0	100
FI	Count	**220**	0	0	0	0	0	0	0	304	0	0	0	0	0	0	0	0	0	0	6	0	0	0	0	530
	%	**42**	0	0	0	0	0	0	0	57	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	100
FR	Count	**464**	0	0	0	0	144	336	0	0	0	201	0	0	0	0	0	0	0	0	0	0	0	0	0	1145
	%	**41**	0	0	0	0	13	29	0	0	0	18	0	0	0	0	0	0	0	0	0	0	0	0	0	100
HU	Count	0	6	0	0	0	0	0	0	0	0	0	46	3	0	0	0	0	0	86	0	0	52	0	0	193
	%	0	3	0	0	0	0	0	0	0	0	0	24	2	0	0	0	0	0	45	0	0	27	0	0	100
IT	Count	**160**	0	0	0	0	0	0	0	0	0	0	0	0	492	0	0	0	0	0	0	0	0	0	0	652
	%	**25**	0	0	0	0	0	0	0	0	0	0	0	0	75	0	0	0	0	0	0	0	0	0	0	100
LT	Count	0	0	0	0	0	0	0	0	0	115	0	0	0	0	0	0	0	0	0	0	0	0	0	0	115
	%	0	0	0	0	0	0	0	0	0	100	0	0	0	0	0	0	0	0	0	0	0	0	0	0	100
NL	Count	0	0	0	0	0	0	0	0	0	0	0	102	0	0	0	0	0	0	79	0	0	1	0	0	182
	%	0	0	0	0	0	0	0	0	0	0	0	56	0	0	0	0	0	0	44	0	0	1	0	0	100
PL	Count	0	0	0	0	0	0	0	0	0	0	0	0	0	0	916	0	0	0	3	0	0	0	0	0	919
	%	0	0	0	0	0	0	0	0	0	0	0	0	0	0	100	0	0	0	0	0	0	0	0	0	100
PT	Count	0	0	923	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	923
	%	0	0	100	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	100
RO	Count	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	263	0	0	0	263
	%	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	100	0	0	0	100
SE	Count	39	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	566	605
	%	6	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	94	100
UK	Count	0	0	0	1987	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1987
	%	0	0	0	100	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	100
Total	Count	1283	71	923	1987	987	144	336	173	304	115	201	764	789	492	916	130	785	169	344	247	263	174	922	566	13.085
	%	10	1	7	15	8	1	3	1	2	1	2	6	6	4	7	1	6	1	3	2	2	1	7	4	100

Table 56: Table: Distribution of eutrophication

Eutrophication							
Country		Missing	No	Low	Intermediate	Extreme	Total
AT	Count	0.0	763.0	137.0	22.0	0.0	922.0
	%	0.0	82.8	14.9	2.4	0.0	100.0
CH	Count	**638.0**	79.0	0.0	0.0	0.0	717.0
	%	**89.0**	11.0	0.0	0.0	0.0	100.0
DE	Count	0.0	39.0	742.0	4.0	0.0	785.0
	%	0.0	5.0	94.5	0.5	0.0	100.0
ES	Count	48.0	2250.0	539.0	267.0	43.0	3147.0
	%	1.5	71.5	17.1	8.5	1.4	100.0
FI	Count	**220.0**	163.0	130.0	13.0	4.0	530.0
	%	**41.5**	30.8	24.5	2.5	0.8	100.0
FR	Count	52.0	532.0	288.0	260.0	13.0	1145.0
	%	4.5	46.5	25.2	22.7	1.1	100.0
HU	Count	0.0	150.0	0.0	43.0	0.0	193.0
	%	0.0	77.7	0.0	22.3	0.0	100.0
IT	Count	24.0	535.0	36.0	31.0	26.0	652.0
	%	3.7	82.1	5.5	4.8	4.0	100.0
LT	Count	0.0	66.0	27.0	19.0	3.0	115.0
	%	0.0	57.4	23.5	16.5	2.6	100.0
NL	Count	0.0	4.0	57.0	88.0	33.0	182.0
	%	0.0	2.2	31.3	48.4	18.1	100.0
PL	Count	2.0	585.0	112.0	191.0	29.0	919.0
	%	0.2	63.7	12.2	20.8	3.2	100.0
PT	Count	0.0	182.0	439.0	248.0	54.0	923.0
	%	0.0	19.7	47.6	26.9	5.9	100.0
RO	Count	0.0	244.0	19.0	0.0	0.0	263.0
	%	0.0	92.8	7.2	0.0	0.0	100.0
SE	Count	0.0	325.0	139.0	141.0	0.0	605.0
	%	0.0	53.7	23.0	23.3	0.0	100.0
UK	Count	7.0	1219.0	39.0	722.0	0.0	1987.0
	%	0.4	61.3	2.0	36.3	0.0	100.0
Total	Count	991.0	7136.0	2704.0	2049.0	205.0	13,085.0
	%	7.6	54.5	20.7	15.7	1.6	100.0

Table 57 Distribution of organic_pollution

Organic_pollution						
Country		Missing	No	Weak	Strong	Total
AT	Count	6.0	312.0	598.0	6.0	922.0
	%	0.7	33.8	64.9	0.7	100.0
CH	Count	**476.0**	241.0	0.0	0.0	717.0
	%	**66.4**	33.6	0.0	0.0	100.0
DE	Count	0.0	39.0	742.0	4.0	785.0
	%	0.0	5.0	94.5	0.5	100.0
ES	Count	93.0	2321.0	709.0	24.0	3147.0
	%	3.0	73.8	22.5	0.8	100.0
FI	Count	**220.0**	310.0	0.0	0.0	530.0
	%	**41.5**	58.5	0.0	0.0	100.0
FR	Count	**245.0**	343.0	491.0	66.0	1145.0
	%	**21.4**	30.0	42.9	5.8	100.0
HU	Count	0.0	93.0	77.0	23.0	193.0
	%	0.0	48.2	39.9	11.9	100.0
IT	Count	74.0	542.0	29.0	7.0	652.0
	%	11.3	83.1	4.4	1.1	100.0
LT	Count	0.0	53.0	50.0	12.0	115.0
	%	0.0	46.1	43.5	10.4	100.0
NL	Count	0.0	34.0	59.0	89.0	182.0
	%	0.0	18.7	32.4	48.9	100.0
PL	Count	2.0	666.0	189.0	62.0	919.0
	%	0.2	72.5	20.6	6.7	100.0
PT	Count	0.0	527.0	326.0	70.0	923.0
	%	0.0	57.1	35.3	7.6	100.0
RO	Count	0.0	210.0	46.0	7.0	263.0
	%	0.0	79.8	17.5	2.7	100.0
SE	Count	0.0	605.0	0.0	0.0	605.0
	%	0.0	100.0	0.0	0.0	100.0
UK	Count	7.0	1219.0	108.0	653.0	1987.0
	%	0.4	61.3	5.4	32.9	100.0
Total	Count	1123.0	7515.0	3424.0	1023.0	13,085.0
	%	8.6	57.4	26.2	7.8	100.0

Table 58: Distribution of organic_siltation

Organic_siltation					
Country		Missing	No	Yes	Total
AT	Count	0.0	922.0	0.0	922.0
	%	0.0	100.0	0.0	100.0
CH	Count	**476.0**	241.0	0.0	717.0
	%	**66.4**	33.6	0.0	100.0
DE	Count	0.0	109.0	676.0	785.0
	%	0.0	13.9	86.1	100.0
ES	Count	35.0	3031.0	81.0	3147.0
	%	1.1	96.3	2.6	100.0
FI	Count	**220.0**	98.0	212.0	530.0
	%	**41.5**	18.5	40.0	100.0
FR	Count	59.0	865.0	221.0	1145.0
	%	5.2	75.5	19.3	100.0
HU	Count	0.0	94.0	99.0	193.0
	%	0.0	48.7	51.3	100.0
IT	Count	**130.0**	487.0	35.0	652.0
	%	**19.9**	74.7	5.4	100.0
LT	Count	1.0	62.0	52.0	115.0
	%	0.9	53.9	45.2	100.0
NL	Count	0.0	43.0	139.0	182.0
	%	0.0	23.6	76.4	100.0
PL	Count	2.0	813.0	104.0	919.0
	%	0.2	88.5	11.3	100.0
PT	Count	0.0	639.0	284.0	923.0
	%	0.0	69.2	30.8	100.0
RO	Count	0.0	227.0	36.0	263.0
	%	0.0	86.3	13.7	100.0
SE	Count	0.0	605.0	0.0	605.0
	%	0.0	100.0	0.0	100.0
UK	Count	**1987.0**	0.0	0.0	1987.0
	%	**100.0**	0.0	0.0	100.0
Total	Count	2910.0	8236.0	1939.0	13,085.0
	%	22.2	62.9	14.8	100.0

Table 59: Distribution of acidification

Acidification					
Country		Missing	No	Yes	Total
AT	Count	0.0	922.0	0.0	922.0
	%	0.0	100.0	0.0	100.0
CH	Count	0.0	717.0	0.0	717.0
	%	0.0	100.0	0.0	100.0
DE	Count	0.0	783.0	2.0	785.0
	%	0.0	99.7	0.3	100.0
ES	Count	0.0	3116.0	31.0	3147.0
	%	0.0	99.0	1.0	100.0
FI	Count	**220.0**	306.0	4.0	530.0
	%	**41.5**	57.7	0.8	100.0
FR	Count	86.0	1043.0	16.0	1145.0
	%	7.5	91.1	1.4	100.0
HU	Count	0.0	192.0	1.0	193.0
	%	0.0	99.5	0.5	100.0
IT	Count	0.0	652.0	0.0	652.0
	%	0.0	100.0	0.0	100.0
LT	Count	0.0	115.0	0.0	115.0
	%	0.0	100.0	0.0	100.0
NL	Count	0.0	182.0	0.0	182.0
	%	0.0	100.0	0.0	100.0
PL	Count	4.0	885.0	30.0	919.0
	%	0.4	96.3	3.3	100.0
PT	Count	0.0	917.0	6.0	923.0
	%	0.0	99.3	0.7	100.0
RO	Count	0.0	263.0	0.0	263.0
	%	0.0	100.0	0.0	100.0
SE	Count	0.0	473.0	132.0	605.0
	%	0.0	78.2	21.8	100.0
UK	Count	7.0	1882.0	98.0	1987.0
	%	0.4	94.7	4.9	100.0
Total	Count	317.0	12,448.0	320.0	13,085.0
	%	2.4	95.1	2.4	100.0

Table 60: Distribution of toxic_substances

Toxic_substances						
Country		Missing	No	Intermediate	High concentration	Total
AT	Count	0.0	852.0	17.0	53.0	922.0
	%	0.0	92.4	1.8	5.7	100.0
CH	Count	**638.0**	79.0	0.0	0.0	717.0
	%	**89.0**	11.0	0.0	0.0	100.0
DE	Count	23.0	111.0	423.0	228.0	785.0
	%	2.9	14.1	53.9	29.0	100.0
ES	Count	18.0	2891.0	203.0	35.0	3147.0
	%	0.6	91.9	6.5	1.1	100.0
FI	Count	**220.0**	310.0	0.0	0.0	530.0
	%	**41.5**	58.5	0.0	0.0	100.0
FR	Count	**253.0**	475.0	336.0	81.0	1145.0
	%	**22.1**	41.5	29.3	7.1	100.0
HU	Count	0.0	193.0	0.0	0.0	193.0
	%	0.0	100.0	0.0	0.0	100.0
IT	Count	**268.0**	384.0	0.0	0.0	652.0
	%	**41.1**	58.9	0.0	0.0	100.0
LT	Count	0.0	108.0	6.0	1.0	115.0
	%	0.0	93.9	5.2	0.9	100.0
NL	Count	0.0	0.0	146.0	36.0	182.0
	%	0.0	0.0	80.2	19.8	100.0
PL	Count	5.0	838.0	70.0	6.0	919.0
	%	0.5	91.2	7.6	0.7	100.0
PT	Count	0.0	725.0	140.0	58.0	923.0
	%	0.0	78.5	15.2	6.3	100.0
RO	Count	0.0	231.0	24.0	8.0	263.0
	%	0.0	87.8	9.1	3.0	100.0
SE	Count	0.0	601.0	4.0	0.0	605.0
	%	0.0	99.3	0.7	0.0	100.0
UK	Count	7.0	160.0	1150.0	670.0	1987.0
	%	0.4	8.1	57.9	33.7	100.0
Total	Count	1432.0	7958.0	2519.0	1176.0	13,085.0
	%	10.9	60.8	19.3	9.0	100.0

Morphological Pressures

Table 61. Distribution of channelisation

Channelisation						
Country		Missing	No	Intermediate	Straightened	Total
AT	Count	0.0	500.0	196.0	226.0	922.0
	%	0.0	54.2	21.3	24.5	100.0
CH	Count	0.0	319.0	96.0	302.0	717.0
	%	0.0	44.5	13.4	42.1	100.0
DE	Count	0.0	85.0	306.0	394.0	785.0
	%	0.0	10.8	39.0	50.2	100.0
ES	Count	263.0	2481.0	291.0	112.0	3147.0
	%	8.4	78.8	9.2	3.6	100.0
FI	Count	**249.0**	265.0	16.0	0.0	530.0
	%	**47.0**	50.0	3.0	0.0	100.0
FR	Count	29.0	815.0	144.0	157.0	1145.0
	%	2.5	71.2	12.6	13.7	100.0
HU	Count	0.0	91.0	70.0	32.0	193.0
	%	0.0	47.2	36.3	16.6	100.0
IT	Count	106.0	389.0	70.0	87.0	652.0
	%	16.3	59.7	10.7	13.3	100.0
LT	Count	0.0	97.0	6.0	12.0	115.0
	%	0.0	84.3	5.2	10.4	100.0
NL	Count	0.0	27.0	49.0	106.0	182.0
	%	0.0	14.8	26.9	58.2	100.0
PL	Count	0.0	579.0	245.0	95.0	919.0
	%	0.0	63.0	26.7	10.3	100.0
PT	Count	0.0	851.0	47.0	25.0	923.0
	%	0.0	92.2	5.1	2.7	100.0
RO	Count	0.0	229.0	9.0	25.0	263.0
	%	0.0	87.1	3.4	9.5	100.0
SE	Count	0.0	492.0	110.0	3.0	605.0
	%	0.0	81.3	18.2	0.5	100.0
UK	Count	**342.0**	1376.0	132.0	137.0	1987.0
	%	**17.2**	69.3	6.6	6.9	100.0
Total	Count	989.0	8596.0	1787.0	1713.0	13,085.0
	%	7.6	65.7	13.7	13.1	100.0

Table 62: Distribution of cross_sec

Cross_sec						
Country		Missing	No	Intermediate	Technical crossec/U-profile	Total
AT	Count	0.0	323.0	358.0	241.0	922.0
	%	0.0	35.0	38.8	26.1	100.0
CH	Count	**160.0**	233.0	77.0	247.0	717.0
	%	**22.3**	32.5	10.7	34.4	100.0
DE	Count	0.0	141.0	97.0	547.0	785.0
	%	0.0	18.0	12.4	69.7	100.0
ES	Count	264.0	2570.0	243.0	70.0	3147.0
	%	8.4	81.7	7.7	2.2	100.0
FI	Count	**249.0**	265.0	16.0	0.0	530.0
	%	**47.0**	50.0	3.0	0.0	100.0
FR	Count	34.0	784.0	151.0	176.0	1145.0
	%	3.0	68.5	13.2	15.4	100.0
HU	Count	0.0	110.0	58.0	25.0	193.0
	%	0.0	57.0	30.1	13.0	100.0
IT	Count	**101.0**	401.0	125.0	25.0	652.0
	%	**15.5**	61.5	19.2	3.8	100.0
LT	Count	0.0	97.0	13.0	5.0	115.0
	%	0.0	84.3	11.3	4.3	100.0
NL	Count	0.0	13.0	13.0	156.0	182.0
	%	0.0	7.1	7.1	85.7	100.0
PL	Count	0.0	602.0	285.0	32.0	919.0
	%	0.0	65.5	31.0	3.5	100.0
PT	Count	0.0	728.0	191.0	4.0	923.0
	%	0.0	78.9	20.7	0.4	100.0
RO	Count	0.0	228.0	10.0	25.0	263.0
	%	0.0	86.7	3.8	9.5	100.0
SE	Count	0.0	605.0	0.0	0.0	605.0
	%	0.0	100.0	0.0	0.0	100.0
UK	Count	**304.0**	1025.0	519.0	139.0	1987.0
	%	**15.3**	51.6	26.1	7.0	100.0
Total	Count	1112.0	8125.0	2156.0	1692.0	13,085.0
	%	8.5	62.1	16.5	12.9	100.0

Table 63: Distribution of embankment

Embankment							
Country		Missing	No	Continuous no permeability	Continuous permeable	Local	Total
AT	Count	0.0	105.0	271.0	412.0	134.0	922.0
	%	0.0	11.4	29.4	44.7	14.5	100.0
CH	Count	**165.0**	126.0	119.0	122.0	185.0	717.0
	%	**23.0**	17.6	16.6	17.0	25.8	100.0
DE	Count	0.0	61.0	482.0	153.0	89.0	785.0
	%	0.0	7.8	61.4	19.5	11.3	100.0
ES	Count	263.0	2099.0	59.0	255.0	471.0	3147.0
	%	8.4	66.7	1.9	8.1	15.0	100.0
FI	Count	**220.0**	310.0	0.0	0.0	0.0	530.0
	%	**41.5**	58.5	0.0	0.0	0.0	100.0
FR	Count	30.0	820.0	50.0	86.0	159.0	1145.0
	%	2.6	71.6	4.4	7.5	13.9	100.0
HU	Count	0.0	94.0	41.0	0.0	58.0	193.0
	%	0.0	48.7	21.2	0.0	30.1	100.0
IT	Count	**104.0**	408.0	17.0	44.0	79.0	652.0
	%	**16.0**	62.6	2.6	6.7	12.1	100.0
LT	Count	0.0	114.0	1.0	0.0	0.0	115.0
	%	0.0	99.1	0.9	0.0	0.0	100.0
NL	Count	0.0	0.0	111.0	59.0	12.0	182.0
	%	0.0	0.0	61.0	32.4	6.6	100.0
PL	Count	1.0	712.0	28.0	54.0	124.0	919.0
	%	0.1	77.5	3.0	5.9	13.5	100.0
PT	Count	0.0	769.0	1.0	24.0	129.0	923.0
	%	0.0	83.3	0.1	2.6	14.0	100.0
RO	Count	0.0	230.0	1.0	15.0	17.0	263.0
	%	0.0	87.5	0.4	5.7	6.5	100.0
SE	Count	0.0	605.0	0.0	0.0	0.0	605.0
	%	0.0	100.0	0.0	0.0	0.0	100.0
UK	Count	**593.0**	1255.0	0.0	57.0	82.0	1987.0
	%	**29.8**	63.2	0.0	2.9	4.1	100.0
Total	Count	1376.0	7708.0	1181.0	1281.0	1539.0	13,085.0
	%	10.5	58.9	9.0	9.8	11.8	100.0

Table 64: Distribution of floodprotection

Floodprotection					
Country		Missing	No	Yes	Total
AT	Count	0.0	100.0	822.0	922.0
	%	0.0	10.8	89.2	100.0
CH	Count	0.0	717.0	0.0	717.0
	%	0.0	100.0	0.0	100.0
DE	Count	0.0	148.0	637.0	785.0
	%	0.0	18.9	81.1	100.0
ES	Count	267.0	2388.0	492.0	3147.0
	%	8.5	75.9	15.6	100.0
FI	Count	**220.0**	305.0	5.0	530.0
	%	**41.5**	57.5	0.9	100.0
FR	Count	28.0	996.0	121.0	1145.0
	%	2.4	87.0	10.6	100.0
HU	Count	0.0	137.0	56.0	193.0
	%	0.0	71.0	29.0	100.0
IT	Count	**108.0**	455.0	89.0	652.0
	%	**16.6**	69.8	13.7	100.0
LT	Count	0.0	114.0	1.0	115.0
	%	0.0	99.1	0.9	100.0
NL	Count	0.0	0.0	182.0	182.0
	%	0.0	0.0	100.0	100.0
PL	Count	3.0	842.0	74.0	919.0
	%	0.3	91.6	8.1	100.0
PT	Count	0.0	880.0	43.0	923.0
	%	0.0	95.3	4.7	100.0
RO	Count	0.0	227.0	36.0	263.0
	%	0.0	86.3	13.7	100.0
SE	Count	0.0	605.0	0.0	605.0
	%	0.0	100.0	0.0	100.0
UK	Count	**604.0**	1219.0	164.0	1987.0
	%	**30.4**	61.3	8.3	100.0
Total	Count	1230.0	9133.0	2722.0	13,085.0
	%	9.4	69.8	20.8	100.0

Table 65: Distribution of floodplain

Floodplain								
Country		Missing	No	Small	Medium	Large	Some water	Total
AT	Count	0.0	825.0	30.0	54.0	13.0	0.0	922.0
	%	0.0	89.5	3.3	5.9	1.4	0.0	100.0
CH	Count	**401.0**	254.0	2.0	41.0	19.0	0.0	717.0
	%	**55.9**	35.4	0.3	5.7	2.6	0.0	100.0
DE	Count	52.0	627.0	35.0	1.0	70.0	0.0	785.0
	%	6.6	79.9	4.5	0.1	8.9	0.0	100.0
ES	Count	**835.0**	2033.0	185.0	74.0	20.0	0.0	3147.0
	%	**26.5**	64.6	5.9	2.4	0.6	0.0	100.0
FI	Count	**220.0**	306.0	2.0	2.0	0.0	0.0	530.0
	%	**41.5**	57.7	0.4	0.4	0.0	0.0	100.0
FR	Count	65.0	715.0	162.0	73.0	65.0	65.0	1145.0
	%	5.7	62.4	14.1	6.4	5.7	5.7	100.0
HU	Count	0.0	135.0	39.0	15.0	4.0	0.0	193.0
	%	0.0	69.9	20.2	7.8	2.1	0.0	100.0
IT	Count	0.0	641.0	7.0	2.0	2.0	0.0	652.0
	%	0.0	98.3	1.1	0.3	0.3	0.0	100.0
LT	Count	109.0	1.0	1.0	4.0	0.0	0.0	115.0
	%	94.8	0.9	0.9	3.5	0.0	0.0	100.0
NL	Count	0.0	147.0	11.0	0.0	0.0	24.0	182.0
	%	0.0	80.8	6.0	0.0	0.0	13.2	100.0
PL	Count	**635.0**	64.0	60.0	71.0	89.0	0.0	919.0
	%	**69.1**	7.0	6.5	7.7	9.7	0.0	100.0
PT	Count	0.0	774.0	34.0	51.0	64.0	0.0	923.0
	%	0.0	83.9	3.7	5.5	6.9	0.0	100.0
RO	Count	9.0	0.0	147.0	61.0	46.0	0.0	263.0
	%	3.4	0.0	55.9	23.2	17.5	0.0	100.0
SE	Count	0.0	605.0	0.0	0.0	0.0	0.0	605.0
	%	0.0	100.0	0.0	0.0	0.0	0.0	100.0
UK	Count	**1987.0**	0.0	0.0	0.0	0.0	0.0	1987.0
	%	**100.0**	0.0	0.0	0.0	0.0	0.0	100.0
Total	Count	4313.0	7127.0	715.0	449.0	392.0	89.0	13,085.0
	%	33.0		5.5	3.4	3.0	0.7	100.0

Table 66: Distribution of instream_habitat

Instream_habitat						
Country		Missing	No	Intermediate	High	Total
AT	Count	0.0	319.0	441.0	162.0	922.0
	%	0.0	34.6	47.8	17.6	100.0
CH	Count	**183.0**	147.0	24.0	363.0	717.0
	%	**25.5**	20.5	3.3	50.6	100.0
DE	Count	0.0	152.0	87.0	546.0	785.0
	%	0.0	19.4	11.1	69.6	100.0
ES	Count	266.0	2480.0	290.0	111.0	3147.0
	%	8.5	78.8	9.2	3.5	100.0
FI	Count	**249.0**	265.0	16.0	0.0	530.0
	%	**47.0**	50.0	3.0	0.0	100.0
FR	Count	35.0	777.0	215.0	118.0	1145.0
	%	3.1	67.9	18.8	10.3	100.0
HU	Count	0.0	40.0	129.0	24.0	193.0
	%	0.0	20.7	66.8	12.4	100.0
IT	Count	**105.0**	378.0	121.0	48.0	652.0
	%	**16.1**	58.0	18.6	7.4	100.0
LT	Count	0.0	106.0	7.0	2.0	115.0
	%	0.0	92.2	6.1	1.7	100.0
NL	Count	0.0	0.0	13.0	169.0	182.0
	%	0.0	0.0	7.1	92.9	100.0
PL	Count	0.0	601.0	270.0	48.0	919.0
	%	0.0	65.4	29.4	5.2	100.0
PT	Count	0.0	628.0	231.0	64.0	923.0
	%	0.0	68.0	25.0	6.9	100.0
RO	Count	0.0	227.0	25.0	11.0	263.0
	%	0.0	86.3	9.5	4.2	100.0
SE	Count	0.0	535.0	70.0	0.0	605.0
	%	0.0	88.4	11.6	0.0	100.0
UK	Count	**304.0**	252.0	1112.0	319.0	1987.0
	%	**15.3**	12.7	56.0	16.1	100.0
Total	Count	1142.0	6907.0	3051.0	1985.0	13,085.0
	%	8.7	52.8	23.3	15.2	100.0

Table 67: Distribution of riparian_vegetation

Riparian_vegetation							
Country		Missing	No	Slight	Intermediate	High	Total
AT	Count	0.0	293.0	201.0	137.0	291.0	922.0
	%	0.0	31.8	21.8	14.9	31.6	100.0
CH	Count	0.0	471.0	30.0	108.0	108.0	717.0
	%	0.0	65.7	4.2	15.1	15.1	100.0
DE	Count	0.0	47.0	73.0	192.0	473.0	785.0
	%	0.0	6.0	9.3	24.5	60.3	100.0
ES	Count	267.0	1156.0	610.0	722.0	392.0	3147.0
	%	8.5	36.7	19.4	22.9	12.5	100.0
FI	Count	**220.0**	310.0	0.0	0.0	0.0	530.0
	%	**41.5**	58.5	0.0	0.0	0.0	100.0
FR	Count	32.0	705.0	194.0	137.0	77.0	1145.0
	%	2.8	61.6	16.9	12.0	6.7	100.0
HU	Count	0.0	23.0	14.0	130.0	26.0	193.0
	%	0.0	11.9	7.3	67.4	13.5	100.0
IT	Count	**105.0**	334.0	56.0	104.0	53.0	652.0
	%	**16.1**	51.2	8.6	16.0	8.1	100.0
LT	Count	0.0	56.0	29.0	18.0	12.0	115.0
	%	0.0	48.7	25.2	15.7	10.4	100.0
NL	Count	0.0	0.0	3.0	10.0	169.0	182.0
	%	0.0	0.0	1.6	5.5	92.9	100.0
PL	Count	1.0	480.0	157.0	133.0	148.0	919.0
	%	0.1	52.2	17.1	14.5	16.1	100.0
PT	Count	0.0	171.0	298.0	321.0	133.0	923.0
	%	0.0	18.5	32.3	34.8	14.4	100.0
RO	Count	0.0	234.0	29.0	0.0	0.0	263.0
	%	0.0	89.0	11.0	0.0	0.0	100.0
SE	Count	0.0	438.0	89.0	57.0	21.0	605.0
	%	0.0	72.4	14.7	9.4	3.5	100.0
UK	Count	**1987.0**	0.0	0.0	0.0	0.0	1987.0
	%	**100.0**	0.0	0.0	0.0	0.0	100.0
Total	Count	2612.0	4718.0	1783.0	2069.0	1903.0	13,085.0
	%	20.0	36.1	13.6	15.8	14.5	100.0

Table 68: Distribution of sedimentation

Sedimentation							
Country		Missing	No	Weak	Medium	High	Total
AT	Count	0.0	68.0	658.0	132.0	64.0	922.0
	%	0.0	7.4	71.4	14.3	6.9	100.0
CH	Count	**635.0**	32.0	0.0	24.0	26.0	717.0
	%	**88.6**	4.5	0.0	3.3	3.6	100.0
DE	Count	8.0	612.0	118.0	39.0	8.0	785.0
	%	1.0	78.0	15.0	5.0	1.0	100.0
ES	Count	**789.0**	1473.0	545.0	246.0	94.0	3147.0
	%	**25.1**	46.8	17.3	7.8	3.0	100.0
FI	Count	**226.0**	32.0	215.0	53.0	4.0	530.0
	%	**42.6**	6.0	40.6	10.0	0.8	100.0
FR	Count	44.0	582.0	251.0	187.0	81.0	1145.0
	%	3.8	50.8	21.9	16.3	7.1	100.0
HU	Count	0.0	149.0	20.0	21.0	3.0	193.0
	%	0.0	77.2	10.4	10.9	1.6	100.0
IT	Count	0.0	633.0	10.0	8.0	1.0	652.0
	%	0.0	97.1	1.5	1.2	0.2	100.0
LT	Count	9.0	93.0	12.0	1.0	0.0	115.0
	%	7.8	80.9	10.4	0.9	0.0	100.0
NL	Count	0.0	175.0	0.0	4.0	3.0	182.0
	%	0.0	96.2	0.0	2.2	1.6	100.0
PL	Count	4.0	701.0	91.0	112.0	11.0	919.0
	%	0.4	76.3	9.9	12.2	1.2	100.0
PT	Count	0.0	337.0	333.0	169.0	84.0	923.0
	%	0.0	36.5	36.1	18.3	9.1	100.0
RO	Count	0.0	252.0	11.0	0.0	0.0	263.0
	%	0.0	95.8	4.2	0.0	0.0	100.0
SE	Count	0.0	377.0	168.0	49.0	11.0	605.0
	%	0.0	62.3	27.8	8.1	1.8	100.0
UK	Count	7.0	923.0	464.0	319.0	274.0	1987.0
	%	0.4	46.5	23.4	16.1	13.8	100.0
Total	Count	1722.0	6439.0	2896.0	1364.0	664.0	13,085.0
	%	13.2	49.2	22.1	10.4	5.1	100.0

Connectivity Pressures

Table 69. Distribution of barriers_river_segment_up

Barriers_river_segment_up						
Country		Missing	No	Partial	Yes	Total
AT	Count	0.0	545.0	5.0	372.0	922.0
	%	0.0	59.1	0.5	40.3	100.0
CH	Count	23.0	256.0	0.0	438.0	717.0
	%	3.2	35.7	0.0	61.1	100.0
DE	Count	4.0	639.0	0.0	142.0	785.0
	%	0.5	81.4	0.0	18.1	100.0
ES	Count	4.0	2721.0	0.0	422.0	3147.0
	%	0.1	86.5	0.0	13.4	100.0
FI	Count	**220.0**	308.0	0.0	2.0	530.0
	%	**41.5**	58.1	0.0	0.4	100.0
FR	Count	70.0	517.0	242.0	316.0	1145.0
	%	6.1	45.2	21.1	27.6	100.0
HU	Count	0.0	158.0	0.0	35.0	193.0
	%	0.0	81.9	0.0	18.1	100.0
IT	Count	0.0	558.0	0.0	94.0	652.0
	%	0.0	85.6	0.0	14.4	100.0
LT	Count	0.0	96.0	0.0	19.0	115.0
	%	0.0	83.5	0.0	16.5	100.0
NL	Count	0.0	125.0	0.0	57.0	182.0
	%	0.0	68.7	0.0	31.3	100.0
PL	Count	0.0	744.0	17.0	158.0	919.0
	%	0.0	81.0	1.8	17.2	100.0
PT	Count	0.0	809.0	26.0	88.0	923.0
	%	0.0	87.6	2.8	9.5	100.0
RO	Count	0.0	241.0	0.0	22.0	263.0
	%	0.0	91.6	0.0	8.4	100.0
SE	Count	0.0	490.0	45.0	70.0	605.0
	%	0.0	81.0	7.4	11.6	100.0
UK	Count	0.0	1597.0	390.0	0.0	1987.0
	%	0.0	80.4	19.6	0.0	100.0
Total	Count	321.0	9804.0	725.0	2235.0	13,085.0
	%	2.5	74.9	5.5	17.1	100.0

Table 70: Distribution of barriers_river_segment_down

Barriers_river_segment_down						
Country		Missing	No	Partial	Yes	Total
AT	Count	0.0	533.0	3.0	386.0	922.0
	%	0.0	57.8	0.3	41.9	100.0
CH	Count	23.0	288.0	0.0	406.0	717.0
	%	3.2	40.2	0.0	56.6	100.0
DE	Count	4.0	609.0	0.0	172.0	785.0
	%	0.5	77.6	0.0	21.9	100.0
ES	Count	15.0	2797.0	0.0	335.0	3147.0
	%	0.5	88.9	0.0	10.6	100.0
FI	Count	**220.0**	309.0	1.0	0.0	530.0
	%	**41.5**	58.3	0.2	0.0	100.0
FR	Count	68.0	553.0	220.0	304.0	1145.0
	%	5.9	48.3	19.2	26.6	100.0
HU	Count	0.0	173.0	0.0	20.0	193.0
	%	0.0	89.6	0.0	10.4	100.0
IT	Count	0.0	552.0	0.0	100.0	652.0
	%	0.0	84.7	0.0	15.3	100.0
LT	Count	0.0	103.0	0.0	12.0	115.0
	%	0.0	89.6	0.0	10.4	100.0
NL	Count	0.0	136.0	0.0	46.0	182.0
	%	0.0	74.7	0.0	25.3	100.0
PL	Count	0.0	737.0	12.0	170.0	919.0
	%	0.0	80.2	1.3	18.5	100.0
PT	Count	0.0	808.0	29.0	86.0	923.0
	%	0.0	87.5	3.1	9.3	100.0
RO	Count	0.0	242.0	0.0	21.0	263.0
	%	0.0	92.0	0.0	8.0	100.0
SE	Count	0.0	558.0	20.0	27.0	605.0
	%	0.0	92.2	3.3	4.5	100.0
UK	Count	0.0	1551.0	436.0	0.0	1987.0
	%	0.0	78.1	21.9	0.0	100.0
Total	Count	330.0	9949.0	721.0	2085.0	13,085.0
	%	2.5	76.0	5.5	15.9	100.0

Table 71: Distribution of barriers_catchment_down

Barriers_catchment_down						
Country		Missing	No	Partial	Yes	Total
AT	Count	0.0	0.0	0.0	922.0	922.0
	%	0.0	0.0	0.0	100.0	100.0
CH	Count	0.0	0.0	0.0	717.0	717.0
	%	0.0	0.0	0.0	100.0	100.0
DE	Count	4.0	79.0	0.0	702.0	785.0
	%	0.5	10.1	0.0	89.4	100.0
ES	Count	0.0	399.0	590.0	2158.0	3147.0
	%	0.0	12.7	18.7	68.6	100.0
FI	Count	**220.0**	218.0	2.0	90.0	530.0
	%	**41.5**	41.1	0.4	17.0	100.0
FR	Count	40.0	133.0	163.0	809.0	1145.0
	%	3.5	11.6	14.2	70.7	100.0
HU	Count	0.0	0.0	0.0	193.0	193.0
	%	0.0	0.0	0.0	100.0	100.0
IT	Count	0.0	253.0	0.0	399.0	652.0
	%	0.0	38.8	0.0	61.2	100.0
LT	Count	0.0	56.0	9.0	50.0	115.0
	%	0.0	48.7	7.8	43.5	100.0
NL	Count	0.0	15.0	0.0	167.0	182.0
	%	0.0	8.2	0.0	91.8	100.0
PL	Count	0.0	140.0	54.0	725.0	919.0
	%	0.0	15.2	5.9	78.9	100.0
PT	Count	0.0	306.0	43.0	574.0	923.0
	%	0.0	33.2	4.7	62.2	100.0
RO	Count	0.0	236.0	0.0	27.0	263.0
	%	0.0	89.7	0.0	10.3	100.0
SE	Count	0.0	205.0	32.0	368.0	605.0
	%	0.0	33.9	5.3	60.8	100.0
UK	Count	0.0	247.0	1026.0	714.0	1987.0
	%	0.0	12.4	51.6	35.9	100.0
Total	Count	264.0	2287.0	1919.0	8615.0	13,085.0
	%	2.0	17.5	14.7	65.8	100.0